

Qy	TTCTCAGCAAACTCCAGCGCTGTATGAGAAACATTAATTCGTTGTTCAGGACAGACTA	60
Db	1 TTCTCAGCAAACTCCAGCGCTGTATGAGAAACATTAATTCGTTGTTCAGGACAGACTA	60
Qy	61 GCACGCTGTGCTCTTCGAGCTTGGCTGCTCGTTCTGAGCTGCACCACTTAAGGTTCAAG	120
Db	61 GCACGCTGTGCTCTTCGAGCTTGGCTGCTCGTTCTGAGCTGCACCACTTAAGGTTCAAG	120
Qy	121 CAACCTCCACGCTTTCTCTTCGCGCTTGGCGTCACTTTCTAAGAAATTCACAGAGGCA	180
Db	121 CAACCTCCACGCTTTCTCTTCGCGCTTGGCGTCACTTTCTAAGAAATTCACAGAGGCA	180

QY 181 GCGAGAGCGGGGCGGGCTCTGAGACTCCGGGCTCCGCTCTTTCCGGGAAACCGCCACTA 240
DB 181 GCGAGAGCGGGGCGGGCTCTGAGACTCCGGGCTCCGCTCTTTCCGGGAAACCGCCACTA 240
QY 241 CCCAGAGCTCCGACAGAGGGTGAAGAAAAGATTAATCTTCGGGTCTCCGATCGTCTTAATC 300
DB 241 CCCAGAGCTCCGACAGAGGGTGAAGAAAAGATTAATCTTCGGGTCTCCGATCGTCTTAATC 300
QY 301 TCGCAGAGAGAGAGCGCGGCCCATCCGGCCGAAACGAGGGGGGAGGGGAGGGGGTG 360
DB 301 TCGCAGAGAGAGAGCGCGGCCCATCCGGCCGAAACGAGGGGGGAGGGGAGGGGGTG 360
QY 361 TGGCCGGGAGCGCGAAATCCCGGAGTAAAGAGAGGGGCGGGGTCCGCGTCCCGG 420
DB 361 TGGCCGGGAGCGCGAAATCCCGGAGTAAAGAGAGGGGCGGGGTCCGCGTCCCGG 420
QY 421 GCATACGATCGTGCACGCTGCGGCTGGGCTGGGCTGAGAGGGGAGGGGGCGGGCG 480
DB 421 GCATACGATCGTGCACGCTGCGGCTGGGCTGAGAGGGGAGGGGGCGGGCG 480
QY 481 GCCGAGCGCGGTGTTATTTCCGTGTCCGACAGTCCGTGCGGGCGCGGGTGCACCG 540
DB 481 GCCGAGCGCGGTGTTATTTCCGTGTCCGACAGTCCGTGCGGGCGCGGGTGCACCG 540
QY 541 GGAGAGTGGCATTAAGTGTATGAAGAGCTCTGTATGATGATGACATTCAGATGGAG 600
DB 541 GGAGAGTGGCATTAAGTGTATGAAGAGCTCTGTATGATGATGACATTCAGATGGAG 600
QY 601 CTCAGTATGCCAGAAAAAATGAGAAAAAGCAATCAAACTGGGTGACATTAACCAAGAT 660
DB 601 CTCAGTATGCCAGAAAAAATGAGAAAAAGCAATCAAACTGGGTGACATTAACCAAGAT 660
QY 661 TTGAAGAGCTTTCGAGATTAAGTGGGAGAACTACTCTTAATGATGATGATGATGATG 720
DB 661 TTGAAGAGCTTTCGAGATTAAGTGGGAGAACTACTCTTAATGATGATGATGATGATG 720
QY 721 CTTTGAAGCATGTCTGTATTTGAATGATGATCCCAAGATGATGATGATGATGATG 780
DB 721 CTTTGAAGCATGTCTGTATTTGAATGATGATCCCAAGATGATGATGATGATGATGATG 780
QY 781 GGAAGCAAGTTATCGAAGAGTCTCAATTTTGAAGAGCTATCAAGATGCGACTATT 840
DB 781 GGAAGCAAGTTATCGAAGAGTCTCAATTTTGAAGAGCTATCAAGATGCGACTATT 840
QY 841 AAAATTAAAGATCTCACCTTGCTGAACCTGATAGGAGTTATGATACATTTTTCGTG 900
DB 841 AAAATTAAAGATCTCACCTTGCTGAACCTGATAGGAGTTATGATACATTTTTCGTG 900
QY 901 TTGATTAAGCTGTGTAAGAGGCAATTCCTGACAGACAGTATTTTACGCTTTTACAT 960
DB 901 TTGATTAAGCTGTGTAAGAGGCAATTCCTGACAGACAGTATTTTACGCTTTTACAT 960
QY 961 CATATTCAGACTTTATAGAGATCCGTATGAGAGGCTTTTCTCGGGAACTTTGAAA 1020
DB 961 CATATTCAGACTTTATAGAGATCCGTATGAGAGGCTTTTCTCGGGAACTTTGAAA 1020
QY 1021 ATCTGTGACATTCGAGAGGAGAAAAGTAAATTAAGCTGCTTTTGAAGAGAGATTTT 1080
DB 1021 ATCTGTGACATTCGAGAGGAGAAAAGTAAATTAAGCTGCTTTTGAAGAGAGATTTT 1080
QY 1081 CAGTCAATGACTTATGATTTAAATGCTAACAGTGTGACAGATCTTGGAGTTACAGGC 1140
DB 1081 CAGTCAATGACTTATGATTTAAATGCTAACAGTGTGACAGATCTTGGAGTTACAGGC 1140
QY 1141 ATGCTTAAAGATGTGAGAGTGCATGCAAGAGAGTAAAGAGTCTCGAAGTGCACA 1200
DB 1141 ATGCTTAAAGATGTGAGAGTGCATGCAAGAGAGTAAAGAGTCTCGAAGTGCACA 1200
QY 1201 GGAGAGAGAGAGATCAGAGAGTGAAGTGAAGACCAACATGTTTACAGATATTCAGC 1260
DB 1201 GGAGAGAGAGAGATCAGAGAGTGAAGTGAAGACCAACATGTTTACAGATATTCAGC 1260

QY 1261 AGAGTGAATTTACTGCTGTGTACTGACAGAGTCTTATAGCCTTTACTAAGAAAGAGACC 1320
DB 1261 AGAGTGAATTTACTGCTGTGTACTGACAGAGTCTTATAGCCTTTACTAAGAAAGAGACC 1320
QY 1321 AGTGTGTGTCAGAACTCAAAAAATTGATGTCAAGCAGAGAGATCTTCTTGTGCATT 1380
DB 1321 AGTGTGTGTCAGAACTCAAAAAATTGATGTCAAGCAGAGAGATCTTCTTGTGCATT 1380
QY 1381 CATATTCATTTGATGATGATGATCCAGGCCCAAGATGATGATCAAAAGAGATCA 1440
DB 1381 CATATTCATTTGATGATGATGATCCAGGCCCAAGATGATGATCAAAAGAGATCA 1440
QY 1441 ATATGATGGGTTTGAACCCCTTGAACCCAGAGCTCTTCCACTTCCCTCCCA 1500
DB 1441 ATATGATGGGTTTGAACCCCTTGAACCCAGAGCTCTTCCACTTCCCTCCCA 1500
QY 1501 TATGCAAAAATTAATTAAGGAGAAAGATGTAATTTTGCAGATTTAATAGATAGA 1560
DB 1501 TATGCAAAAATTAATTAAGGAGAAAGATGTAATTTTGCAGATTTAATAGATAGA 1560
QY 1561 ATAAAAAGTGTGTGAGGTTGTGAATTTAAGAAATTTAATTTATCTGTGATTTTTC 1620
DB 1561 ATAAAAAGTGTGTGAGGTTGTGAATTTAAGAAATTTAATTTATCTGTGATTTTTC 1620
QY 1621 TGTGAATTTAGTGAACAGTCAACATGTGTTCTTTGAGATCTCTGTTACAAACCACTTC 1680
DB 1621 TGTGAATTTAGTGAACAGTCAACATGTGTTCTTTGAGATCTCTGTTACAAACCACTTC 1680
QY 1681 CTGGTGGATTAACAAAAAGTCTTTTGAAGTCTCTCATGCAAGACATGTTGAAGTGA 1740
DB 1681 CTGGTGGATTAACAAAAAGTCTTTTGAAGTCTCTCATGCAAGACATGTTGAAGTGA 1740
QY 1741 CTTGGTCTTTTGTGAGATCTCGAGTCTTCCGAGAGTCACTATTAATTAATC 1800
DB 1741 CTTGGTCTTTTGTGAGATCTCGAGTCTTCCGAGAGTCACTATTAATTAATC 1800
QY 1801 ACCAGGCTAAGAGCTGTACGACTCTTTTGTACTGACTGTGTGGCCATTCGTAGTC 1860
DB 1801 ACCAGGCTAAGAGCTGTACGACTCTTTTGTACTGACTGTGTGGCCATTCGTAGTC 1860
QY 1861 TTATTCAGATCCATGAGACATAACAGGCTCGACAGAGATTAAGCTGTGATTTCTTG 1920
DB 1861 TTATTCAGATCCATGAGACATAACAGGCTCGACAGAGATTAAGCTGTGATTTCTTG 1920
QY 1921 AGGAATTTGCCACTTTCAGAGATGATTTATGACATTTTATTAATAGCAGAGAGT 1980
DB 1921 AGGAATTTGCCACTTTCAGAGATGATTTATGACATTTTATTAATAGCAGAGAGT 1980
QY 1981 TGATCAGGCTTTCACACGATGCTGTGAACAGGAGCCCAAGGCAACATTTGGCCTG 2040
DB 1981 TGATCAGGCTTTCACACGATGCTGTGAACAGGAGCCCAAGGCAACATTTGGCCTG 2040
QY 2041 TTAGAGTACCTGGGCTTTTACCATTAACCTTGCATTAATGATGATGATGATGATG 2100
DB 2041 TTAGAGTACCTGGGCTTTTACCATTAACCTTGCATTAATGATGATGATGATGATGATG 2100
QY 2101 CTTGAATTTGGAACCTTCAGATGACAGAGATGATGATGATGATGATGATGATGATG 2160
DB 2101 CTTGAATTTGGAACCTTCAGATGACAGAGATGATGATGATGATGATGATGATGATGATG 2160
QY 2161 ATTCCTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
DB 2161 ATTCCTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
QY 2221 GGAAGAGTATGGAAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
DB 2221 GGAAGAGTATGGAAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY 2281 GAAAGTCCGCCATGAGCCAGAGATCAGATGAGCCAGAGATGAGAGATGAGAGATGAG 2340
DB 2281 GAAAGTCCGCCATGAGCCAGAGATCAGATGAGCCAGAGATGAGAGATGAGAGATGAGAG 2340
QY 2341 TGAATGTTTAAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400

```

Db 2341 TGGAACTTTAAACCATGTTACATTTGACATGACGGGAAAGTACGTAACCGAAGTT 2400
Qy 2401 TGAAGTTGATGAAACAAAGTTGGATGAAACAGAGTTTGTCCATTCAACAGTGTAT 2460
Db 2401 TGAAGTTGATGAAACAAAGTTGGATGAAACAGAGTTTGTCCATTCAACAGTGTAT 2460
Qy 2461 GACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2520
Db 2461 GACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2520
Qy 2521 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2580
Db 2521 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2580
Qy 2581 AATGATATGGAATAATATCTTAAACCGGACCTGAGTTAATGAAATTTAAAGTTGC 2640
Db 2581 AATGATATGGAATAATATCTTAAACCGGACCTGAGTTAATGAAATTTAAAGTTGC 2640
Qy 2641 CAACCCCACTTTGTGTTATGAGTTATGAGGAGACCAAAAGAACTTAAAGT 2700
Db 2641 CAACCCCACTTTGTGTTATGAGTTATGAGGAGACCAAAAGAACTTAAAGT 2700
Qy 2701 TCCCTCTGATTTGATTTCTCTGCTCATTAATTTCTCTGTTGAACTTTTGA 2760
Db 2701 TCCCTCTGATTTGATTTCTCTGCTCATTAATTTCTCTGTTGAACTTTTGA 2760
Qy 2761 GAGACTGGGAGGAGTGGCCATAAAGGAGGAGAGTCTTCTTGAACCACTTTAGAGG 2820
Db 2761 GAGACTGGGAGGAGTGGCCATAAAGGAGGAGAGTCTTCTTGAACCACTTTAGAGG 2820
Qy 2821 CACATCACAGGCTCCACATCAGGGAAGTGAAGTATTTCTTGGTACCACTTAA 2880
Db 2821 CACATCACAGGCTCCACATCAGGGAAGTGAAGTATTTCTTGGTACCACTTAA 2880
Qy 2881 TAAGAACTATTTGTTGTTGACAGCTTATATGATGATGAATGAACCTGTTTAAAG 2940
Db 2881 TAAGAACTATTTGTTGTTGACAGCTTATATGATGATGAATGAACCTGTTTAAAG 2940
Qy 2941 TGGTTATTTATGTTCCATGAGAAAGAACTGTTTATGATGATGAATGAACCTTAT 3000
Db 2941 TGGTTATTTATGTTCCATGAGAAAGAACTGTTTATGATGATGAATGAACCTTAT 3000
Qy 3001 GGTATTTATGAGATTTTATGATGATGAATGAATGAATGAATGAATGAATGAATGA 3060
Db 3001 GGTATTTATGAGATTTTATGATGATGAATGAATGAATGAATGAATGAATGAATGA 3060
Qy 3061 TTAATAAGTTAATAATTTCTTGAACAAAAA 3096
Db 3061 TTAATAAGTTAATAATTTCTTGAACAAAAA 3096

```

RESULT 2
US-10-001-857-42.rnpb

Sequence 20725, Application US/09814353
Publication No. US20030165831A1
GENERAL INFORMATION:
APPLICANT: Lee, John
APPLICANT: Thompson, Pamela
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: NMI-006B
CURRENT APPLICATION NUMBER: US/09/814,353
PRIOR FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-25
PRIOR FILING DATE: 2000-05-25
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 60/211,940
PRIOR APPLICATION NUMBER: US 60/216,820

disclose
Appl do not
Prior Art

PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US 60/220,661
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: US 60/257,672
NUMBER OF SEQ ID NOS: 22037
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 20725
LENGTH: 2610
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1, 2, 3, 4
OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20725

Query Match 72.0%; Score 2228.4; DB 12; Length 2610;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 2299; Conservative 0; Mismatches 21; Indels 28; Gaps 3;

```

Qy 426 CGCATGCGTGCAGCGCTGCGGCTGGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 485
Db 2610 CGCATGCGTGCAGCGCTGCGGCTGGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2551
Qy 486 GCGGCGCTGTTATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 545
Db 2550 GCGGCGCTGTTATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2491
Qy 546 AGTAGGATATGTTATGAAAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 605
Db 2490 AGTAGGATATGTTATGAAAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2431
Qy 606 TATGCCAGAAAAATGAGAAAAAGCAATACAACTGGGTGGAATTAACCAAGATTTTGA 665
Db 2430 TATGCCAGAAAAATGAGAAAAAGCAATACAACTGGGTGGAATTAACCAAGATTTTGA 2371
Qy 666 AGAAGCTTGTGGAATTAAGTTGGAGACTATTCATGATGATGATGATGATGATGATGATGAT 725
Db 2370 AGAAGCTTGTGGAATTAAGTTGGAGACTATTCATGATGATGATGATGATGATGATGATGAT 2311
Qy 726 TGAAGCATGCTGCTGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
Db 2310 TGAAGCATGCTGCTGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2251
Qy 786 CCAAGTTATCGAAAAAGTTCTCAATTTTGAACAGCTATCAAGATGAGACTTAAAT 845
Db 2250 CCAAGTTATCGAAAAAGTTCTCAATTTTGAACAGCTATCAAGATGAGACTTAAAT 2191
Qy 846 TAAAGATCTACCTGCTGAACTGATAGGATTAATGATGATGATGATGATGATGATGATGAT 905
Db 2190 TAAAGATCTACCTGCTGAACTGATAGGATTAATGATGATGATGATGATGATGATGATGAT 2131
Qy 906 AACGTGTTAAGAGGCAATCACTGACAGACAGATTTTACGTCCTTACATTA 965
Db 2130 AACGTGTTAAGAGGCAATCACTGACAGACAGATTTTACGTCCTTACATTA 2071
Qy 966 TCCAGCTTTATAGAAAGATCCGCTATGAAGCTTTTGTCTGCGAATTTTGA 1025
Db 2070 TCCAGCTTTATAGAAAGATCCGCTATGAAGCTTTTGTCTGCGAATTTTGA 2011
Qy 1026 TCACATTCAGAGAAAAAGTAAATAAGCTGCTTTTGAAGAGAGAGATTTTCA 1085
Db 2010 TCACATTCAGAGAAAAAGTAAATAAGCTGCTTTTGAAGAGAGAGATTTTCA 1051
Qy 1086 AATGATATGATTTAAATAGGCTTAACAGTGTGACAGATCTTGAAGTGAAGGAGATGCT 1145
Db 1950 AATGATATGATTTAAATAGGCTTAACAGTGTGACAGATCTTGAAGTGAAGGAGATGCT 1891
Qy 1146 AAAAGATGAGAGTACATGCAAGAGAGATTAAGATGCTGAAGTGAAGAGAGA 1205
Db 1890 AAAAGATGAGAGTACATGCAAGAGAGATTAAGATGCTGAAGTGAAGAGAGA 1831

```

QY 1206 AGAAGAGATCCAGAACTGTAAGTGAACCAACAACTGTTAGCAGTATTCAGCAGGT 1265
 Db 1830 AGAAGAGATCCAGAACTGTAAGTGAACCAACAACTGTTAGCAGTATTCAGCAGGT 1771
 QY 1266 GAAATTTACTCGTGTGTGTACTGACAGTGTCTTATAGCCTTACTAAGAAAGACAGTGC 1325
 Db 1770 GAAATTTACTCGTGTGTGTACTGACAGTGTCTTATAGCCTTACTAAGAAAGACAGTGC 1711
 QY 1326 TGTTCGCAAGCTCAAAAATTGATGTTCAAGCAGCAGATCTTCTTCCGCAATCATTA 1385
 Db 1710 TGTTCGCAAGCTCAAAAATTGATGTTCAAGCAGCAGATCTTCTTCCGCAATCATTA 1651
 QY 1386 TTATTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1445
 Db 1650 TTATTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591
 QY 1446 GATGGGTTTGAACCCCTTGTGAACCAAGGCTACTTCCACTTCCCTCGATATGC 1505
 Db 1590 GATGGGTTTGAACCCCTTGTGAACCAAGGCTACTTCCACTTCCCTCGATATGC 1531
 QY 1506 AAAAATAATTAAGGAAAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1565
 Db 1530 AAAAATAATTAAGGAAAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1471
 QY 1566 AACTGTCTGTAGGTTGTGAATTTAACTGATGATGATGATGATGATGATGATGATGAT 1625
 Db 1470 AACTGTCTGTAGGTTGTGAATTTAACTGATGATGATGATGATGATGATGATGATGAT 1411
 QY 1626 AATTAGTGAACAGTCAAGTGTGTTCTTCAAGTGTGTTCTGTTCAACCACTTCTGTGT 1685
 Db 1410 AATTAGTGAACAGTCAAGTGTGTTCTTCAAGTGTGTTCTGTTCAACCACTTCTGTGT 1351
 QY 1686 GATTAACAAAAGTCTTTGGAACCTCATCTCATGCAAGACATGATGAAAGATGCACTGC 1745
 Db 1350 GATTAACAAAAGTCTTTGGAACCTCATCTCATGCAAGACATGATGAAAGATGCACTGC 1291
 QY 1746 GTCTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1805
 Db 1290 GTCTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1233
 QY 1806 GCTAAGAGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1865
 Db 1232 GCTAAGAGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
 QY 1866 CAGATTCATGACATTAACAGGCTGACAGAGACATTAAGCTTGATATTTCTTGAGGAA 1925
 Db 1172 CAGATTCATGACATTAACAGGCTGACAGAGACATTAAGCTTGATATTTCTTGAGGAA 1113
 QY 1926 TTGCGCACCCTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1985
 Db 1112 TTGCGCACCCTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
 QY 1986 CAGGCTTCAACAGCTGTGTTGAAACAGAAACCCCAAGCAACATTTGCGCTGTAG 2045
 Db 1075 CAGGCTTCAACAGCTGTGTTGAAACAGAAACCCCAAGCAACATTTGCGCTGTAG 1016
 QY 2046 GTACCTGGCTCTTTACCATTAACCTTGCATTAATGATGATGATGATGATGATGATGAT 2105
 Db 1015 GTACCTGGCTCTTTACCATTAACCTTGCATTAATGATGATGATGATGATGATGATGAT 956
 QY 2106 AATTGGAACCTTACAGTATGACAGAGTCTATTACATATATTTGATGATGATGATGAT 2165
 Db 955 AATTGGAACCTTACAGTATGACAGAGTCTATTACATATATTTGATGATGATGATGAT 896
 QY 2166 TTTAAGCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2225
 Db 895 TTTAAGCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836
 QY 2226 GGATTAATGGAACAGACAGCAAGAGCGGTGATGATGATGATGATGATGATGATGATGAT 2285
 Db 835 GGATTAATGGAACAGACAGCAAGAGCGGTGATGATGATGATGATGATGATGATGATGAT 776
 QY 2286 TTGCGCCATTTGAGCCGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2345

Db 775 TTGCGCCATTTGAGCCGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 716
 QY 2346 TGTTTAAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2405
 Db 715 TGTTTAAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 656
 QY 2406 TTGATAGGAACATTTGGTATGTAACAGAGTTTGTCTCATTTCAACAGTGTGATGACC 2465
 Db 655 TTGATAGGAACATTTGGTATGTAACAGAGTTTGTCTCATTTCAACAGTGTGATGACC 596
 QY 2466 CGCCCGCAGTCACTACTTACAGTTCAGAGAAATGTGACCTCAATATATATGACCTC 2525
 Db 595 CGCCCGCAGTCACTACTTACAGTTCAGAGAAATGTGACCTCAATATATATGACCTC 536
 QY 2526 CTCTCAGTCTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2585
 Db 535 CTCTCAGTCTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 476
 QY 2586 TATTGGAATAATTTCTTAACCCGACATGAGTTAATAGATTTTAAGTTGACCAAC 2645
 Db 475 TATTGGAATAATTTCTTAACCCGACATGAGTTAATAGATTTTAAGTTGACCAAC 416
 QY 2646 CCAACTTTGTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2705
 Db 415 CCAACTTTGTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 356
 QY 2706 CTGAATTTGATTTCTGCTCATTAATATTTCTGTTGGAACCTGTTGAGAGAC 2765
 Db 355 CTGAATTTGATTTCTGCTCATTAATATTTCTGTTGGAACCTGTTGAGAGAC 299
 QY 2766 TGGGAGAG 2773
 Db 298 GGAAGTGG 291

RESULT 3
 US-09-814-353-15927/C
 : Sequence 15927, Application US/09814353
 : Publication No. US20030165831A1
 : GENERAL INFORMATION:
 : APPLICANT: Lee, John
 : APPLICANT: Thompson, Pamela
 : TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 : TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 : FILE OF INVENTION: THERAPY OF OVARIAN CANCER
 : FILE REFERENCE: MRI-006B
 : CURRENT APPLICATION NUMBER: US/09/814,353
 : CURRENT FILING DATE: 2001-03-21
 : PRIOR APPLICATION NUMBER: US 60/191,031
 : PRIOR FILING DATE: 2000-03-21
 : PRIOR APPLICATION NUMBER: US 60/207,124
 : PRIOR FILING DATE: 2000-05-25
 : PRIOR APPLICATION NUMBER: US 60/211,940
 : PRIOR FILING DATE: 2000-06-15
 : PRIOR APPLICATION NUMBER: US 60/216,820
 : PRIOR FILING DATE: 2000-07-07
 : PRIOR APPLICATION NUMBER: US 60/220,661
 : PRIOR FILING DATE: 2000-07-25
 : PRIOR APPLICATION NUMBER: US 60/257,672
 : PRIOR FILING DATE: 2000-12-21
 : NUMBER OF SEQ ID NOS: 22037
 : SOFTWARE: Faast/Seq for Windows Version 4.0
 : SEQ ID NO 15927
 : LENGTH: 745
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: 571
 : OTHER INFORMATION: n = A,T,C or G
 : US-09-814-353-15927

Query Match 18.1%; Score 560.8; DB 12; Length 745;
 Best Local Similarity 94.8%; Pred. No. 2.2e-141;
 Matches 621; Conservative 0; Mismatches 8; Indels 26; Gaps 3;

QY 1399 GGCATCCAGGCCCCGANTGATGATACAGAAAAGAGATATCCATTTATGATGCGTTTGA 1458
 DB 739 GGCATCCAGGCCCCGANTGATGATACAGAAAAGAGATATCCATTTATGATGCGTTTGA 680
 QY 1459 CCCCTTGAGACGAGAGCTACTTCCACCTCCCTCGATATGCAAAAATATATATA 1518
 DB 679 CCCCTTGAGACGAGAGCTACTTCCACCTCCCTCGATATGCAAAAATATATATA 620
 QY 1519 AGGAAAGAAATGGGAACTATTTTGCAGTTAATAGATAGATATAAAATCTGCTGAG 1578
 DB 619 AGGAAAGAAATGGGAACTATTTTGCAGTTAATAGATAGATATAAAATCTGCTGAG 560
 QY 1579 GTTGTGAATTTACAAATTTACATTTGATCTGGA-TTTTCTGTGAATTTAGTGA 1637
 DB 559 GTTGTGAATTTACAAATTTACATTTGATCTGGA-TTTTCTGTGAATTTAGTGA 500
 QY 1638 GTACCATGTTCTTTCAAGATCTCTTTACAAACCACTTCTGCTGTGATTAACAAAA 1697
 DB 499 GTACCATGTTCTTTCAAGATCTCTTTACAAACCACTTCTGCTGTGATTAACAAAA 440
 QY 1698 GGTCTTGGAACTCATCTCATGACAGATGAGTGAAGATGACCTCGCTTTTGTGAG 1757
 DB 439 GGTCTTGGAACTCATCTCATGACAGATGAGTGAAGATGACCTCGCTTTTGTGAG 380
 QY 1758 ATCTCTCGAGTCTTTCCCAAGTGTACTATATATATATATACAGAGCTAAGACTGT 1817
 DB 379 ATCTCTCGAGTCTTTCCCAAGTGTACTATATATATATATACAGAGCTAAGACTGT 322
 QY 1818 ATGACATCTCTTGTACTCATCTGTTGGCCATCTGATCTTATTTAGATCCATGGA 1877
 DB 321 ATGACATCTCTTGTACTCATCTGTTGGCCATCTGATCTTATTTAGATCCATGGA 262
 QY 1878 CATACAGGAGCTGACAGAGATTAAGCTTGTATCTTGAAGATTTCCACCTTG 1937
 DB 261 CATACAGGAGCTGACAGAGATTAAGCTTGTATCTTGAAGATTTCCACCTTG 202
 QY 1938 CAGAGTGAATTTATGATTTTATTTATAGGACAGAGAGTTGATGACGCTTCA 1997
 DB 201 CAGAGTGAATTTATGATTTTATTTATAGGACAGAGAGTTGATGACGCTTCA 165
 QY 1998 CCATGCTGTTGAACGAGAACCCCAAGGCAATTTGCTGTTAGTACTTG 2052
 DB 164 CCATGCTGTTGAACGAGAACCCCAAGGCAATTTGCTGTTAGTACTTG 110

RESULT 4
 US-09-814-353-16155
 ; Sequence 16155, Application US/09814353
 ; Publication No. US20030165831A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thompson, Pamela
 ; APPLICANT: Lee, John
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; FILE REFERENCE: MRI-006B
 ; CURRENT APPLICATION NUMBER: US/09/814,353
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/191,031
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: US 60/207,124
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: US 60/211,940
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: US 60/216,820
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: US 60/220,661

; PRIOR FILING DATE: 2000-07-25
 ; PRIOR APPLICATION NUMBER: US 60/257,672
 ; PRIOR FILING DATE: 2000-12-21
 ; NUMBER OF SEQ ID NOS: 22037
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 16155
 ; LENGTH: 610
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-814-353-16155

Query Match 16.5%; Score 509.4; DB 12; Length 610;
 Best Local Similarity 99.8%; Pred. No. 1.8e-127;
 Matches 510; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1184 GTACTCGAAGTGAAGAGAGAGAGAGAGATCCGAAGTTGAATAGAACCAACAAT 1243
 DB 100 GTACTCGAAGTGAAGAGAGAGAGAGAGATCCGAAGTTGAATAGAACCAACAAT 159
 QY 1244 GTTACGAGTATTCAGAGAGTGAATTTACTCGTGTGTTACTGACAGTGTATAGCT 1303
 DB 160 GTTACGAGTATTCAGAGAGTGAATTTACTCGTGTGTTACTGACAGTGTATAGCT 219
 QY 1304 TTACTAAGAAAGACCCAGTCTGTTGAGAGCTCAAAAATGATGTTCAAGCAGAG 1363
 DB 220 TTACTAAGAAAGACCCAGTCTGTTGAGAGCTCAAAAATGATGTTCAAGCAGAG 279
 QY 1364 ATCTCTTGTGCAATTCATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1423
 DB 280 ATCTCTTGTGCAATTCATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 339
 QY 1424 CAAAAGAGATCATCAATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1483
 DB 340 CAAAAGAGATCATCAATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 399
 QY 1484 CACCTACCTCCCTCGATATGCAAAAATATATTAAGGAGAAATGTTGAATTTTG 1543
 DB 400 CACCTACCTCCCTCGATATGCAAAAATATATTAAGGAGAAATGTTGAATTTTG 459
 QY 1544 CAAATTAATGATAGATATAAACTGCTGTGAGTTGATGATTAACAATTTTACAT 1603
 DB 460 CAAATTAATGATAGATATAAACTGCTGTGAGTTGATGATTAACAATTTTACAT 519
 QY 1604 GTATCTCGATTTTCTGGAATTTAGTGAACGTCACCTGTTCTTCAAGATCTC 1663
 DB 520 GTATCTCGATTTTCTGGAATTTAGTGAACGTCACCTGTTCTTCAAGATCTC 579
 QY 1664 GTTACCAACCACTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1694
 DB 580 GTTACCAACCACTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 610

RESULT 5
 US-09-814-353-3218/c
 ; Sequence 3218, Application US/09814353
 ; Publication No. US20030165831A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thompson, Pamela
 ; APPLICANT: Lee, John
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; FILE REFERENCE: MRI-006B
 ; CURRENT APPLICATION NUMBER: US/09/814,353
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/191,031
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: US 60/207,124
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: US 60/211,940
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: US 60/216,820

PRIOR FILING DATE: 2000-07-07
 PRIOR APPLICATION NUMBER: US 60/220,661
 PRIOR FILING DATE: 2000-07-25
 PRIOR APPLICATION NUMBER: US 60/257,672
 PRIOR FILING DATE: 2000-12-21
 NUMBER OF SEQ ID NOS: 22037
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3218
 LENGTH: 602
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 561..590
 OTHER INFORMATION: n = A,T,C or G
 US-09-814-353-3218

Query Match 16.0%; Score 496.2; DB 12; Length 602;
 Best Local Similarity 93.9%; Pred. No. 6,88-124;
 Matches 568; Conservative 0; Mismatches 10; Indels 27; Gaps 4;

QY 1450 GGTGTTGACCCCTTGTGAACGAGGCTACTCCACCTACCTCCCTCGATATGC-AAA 1508
 Db 602 GGTGTTGACCCCTTGTGAACGAGGCTACTCCACCTACCTCCCTCGATATGC-AAA 543
 QY 1509 AATTAATTAAGGAGAGAAATGTGAATATTTTGCAGATTAATAGATGAATAAAC 1568
 Db 542 AATTAATTAAGGAGAGAAATGTGAATATTTTGCAGATTAATAGATGAATAAAC 483
 QY 1569 TGTCTGTGAGGTTGTGAATTAATCAATTTACATGATGATCTCTGA-TTTTTCTGTGAAT 1627
 Db 482 TGTCTGTGAGGTTGTGAATTAATCAATTTACATGATGATGATCTCTGA-TTTTTCTGTGAAT 423
 QY 1628 TTAGTGAACAGTCAACCATGTTCTTCAAGATCTCTGTAAACAACCACTTCTCGATG 1687
 Db 422 TTAGTGAACAGTCAACCATGTTCTTCAAGATCTCTGTAAACAACCACTTCTCGATG 363
 QY 1688 AATAACAAAAGGTTGTGAATCTATCTCATGCAAGCATGTGGAAGATGCACTTCGGT 1747
 Db 362 AATAACAAAAGGTTGTGAATCTATCTCATGCAAGCATGTGGAAGATGCACTTCGGT 303
 QY 1748 CTTTGTGAGATCCTCGAGTGTCTTCCCAAGTGTACTATTAATATATACAGG 1807
 Db 302 CTTTGTGAG--TCTTCGGATGCTTCCCAAGTGTACTATTAATATATACAGG 245
 QY 1808 TAAGGACTGTATGACCTCTTTTGTACTCACTGTGTCGCCAATCTGTAGTCTTATGA 1867
 Db 244 TAAGGACTGTATGACCTCTTTTGTACTCACTGTGTCGCCAATCTGTAGTCTTATGA 185
 QY 1868 GATTCATGACATTAACAGGCTCGACAGAGATTAAGCTGTATCTTGTAGGAAT 1927
 Db 184 GATTCATGACATTAACAGGCTCGACAGAGATTAAGCTGTATCTTGTAGGAAT 125
 QY 1928 TGCCACCTTGCAAGATGATTTATGACATTTTATTAATAGCAGAGAGGTTGATGA 1987
 Db 124 TGCCACCTTGCAAGATG-----AGCAGAGAGAGGTTGATGA 88
 QY 1988 GCGCTTACACCATGCTGTGAAACAGAGAACCCCAAGGCAATTTGGCTGTTAGGT 2047
 Db 87 GCGCTTACACCATGCTGTGAAACAGAGAACCCCAAGGCAATTTGGCTGTTAGGT 28
 QY 2048 ACCTG 2052
 Db 27 ACCTG 23

RESULT 6
 US-09-814-353-9543/c
 Sequence 9543, Application US/09814353
 Publication No. US20030165831A1
 GENERAL INFORMATION:
 APPLICANT: Lee, John
 APPLICANT: Thompson, Pamela

APPLICANT: Lillie, James
 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
 FILE REFERENCE: MRI-006B
 CURRENT APPLICATION NUMBER: US/09/814,353
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,031
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/207,124
 PRIOR FILING DATE: 2000-05-25
 PRIOR APPLICATION NUMBER: 60/211,940
 PRIOR FILING DATE: 2000-06-15
 PRIOR APPLICATION NUMBER: 60/216,820
 PRIOR FILING DATE: 2000-07-07
 PRIOR APPLICATION NUMBER: 60/220,661
 PRIOR FILING DATE: 2000-07-25
 PRIOR APPLICATION NUMBER: US 60/257,672
 PRIOR FILING DATE: 2000-12-21
 NUMBER OF SEQ ID NOS: 22037
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 9543
 LENGTH: 602
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 561..590
 OTHER INFORMATION: n = A,T,C or G
 US-09-814-353-9543

Query Match 16.0%; Score 496.2; DB 12; Length 602;
 Best Local Similarity 93.9%; Pred. No. 6,88-124;
 Matches 568; Conservative 0; Mismatches 10; Indels 27; Gaps 4;

QY 1450 GGTGTTGACCCCTTGTGAACGAGGCTACTCCACCTACCTCCCTCGATATGC-AAA 1508
 Db 602 GGTGTTGACCCCTTGTGAACGAGGCTACTCCACCTACCTCCCTCGATATGC-AAA 543
 QY 1509 AATTAATTAAGGAGAGAAATGTGAATATTTTGCAGATTAATAGATGAATAAAC 1568
 Db 542 AATTAATTAAGGAGAGAAATGTGAATATTTTGCAGATTAATAGATGAATAAAC 483
 QY 1569 TGTCTGTGAGGTTGTGAATTAATCAATTTACATGATGATCTCTGA-TTTTTCTGTGAAT 1627
 Db 482 TGTCTGTGAGGTTGTGAATTAATCAATTTACATGATGATCTCTGA-TTTTTCTGTGAAT 423
 QY 1628 TTAGTGAACAGTCAACCATGTTCTTCAAGATCTCTGTAAACAACCACTTCTCGATG 1687
 Db 422 TTAGTGAACAGTCAACCATGTTCTTCAAGATCTCTGTAAACAACCACTTCTCGATG 363
 QY 1688 AATAACAAAAGGTTGTGAATCTATCTCATGCAAGCATGTGGAAGATGCACTTCGGT 1747
 Db 362 AATAACAAAAGGTTGTGAATCTATCTCATGCAAGCATGTGGAAGATGCACTTCGGT 303
 QY 1748 CTTTGTGAGATCCTCGAGTGTCTTCCCAAGTGTACTATTAATATATACAGG 1807
 Db 302 CTTTGTGAG--TCTTCGGATGCTTCCCAAGTGTACTATTAATATATACAGG 245
 QY 1808 TAAGGACTGTATGACCTCTTTTGTACTCACTGTGTCGCCAATCTGTAGTCTTATGA 1867
 Db 244 TAAGGACTGTATGACCTCTTTTGTACTCACTGTGTCGCCAATCTGTAGTCTTATGA 185
 QY 1868 GATTCATGACATTAACAGGCTCGACAGAGATTAAGCTGTATCTTGTAGGAAT 1927
 Db 184 GATTCATGACATTAACAGGCTCGACAGAGATTAAGCTGTATCTTGTAGGAAT 125
 QY 1928 TGCCACCTTGCAAGATGATTTATGACATTTTATTAATAGCAGAGAGGTTGATGA 1987
 Db 124 TGCCACCTTGCAAGATG-----AGCAGAGAGAGGTTGATGA 88
 QY 1988 GCGCTTACACCATGCTGTGAAACAGAGAACCCCAAGGCAATTTGGCTGTTAGGT 2047

Db 87 GCGCTTCACACCATCTGTGTAACAGAACCCCAAGGCAACATTGGCTGTAGT 28
Qy 2048 ACCTG 2052
Db 27 ACCTG 23

RESULT 7
US-09-918-995-20914
; Sequence 20914, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; FROM VARIOUS CDNA LIBRARIES
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20914
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(482)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-20914

Query Match 14.6%; Score 451.2; DB 11; Length 482;
Best Local Similarity 99.3%; Pred.No.9.9e-112;
Matches 453; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 491 CGTCGTATTTCCGTGTCGAGACAGTGGCGCGGCGGAGTACCAAGGAGAGT 550
Db 26 CCAGCTTATTTCCGTGTCGAGACAGTGGCGCGGCGGAGTACCAAGGAGAGT 85
Qy 551 GATATGTTATGAAAGCTTCTGTAGATGAGAGATTACAGATGGAGCTCACTATGC 610
Db 86 GATATGTTATGAAAGCTTCTGTAGATGAGAGATTACAGATGGAGCTCACTATGC 145
Qy 611 CAGAAAAATGAGAAAAACAATACAACTGGGTGGAATCCCAAGATTTTGAAG 670
Db 146 CAGAAAAATGAGAAAAACAATACAACTGGGTGGAATCCCAAGATTTTGAAG 205
Qy 671 CTTCGAGATTAAGTTGGAGAACTTCTCATGATTAAGCTATTGCTTTTGAAG 730
Db 206 CTTCGAGATTAAGTTGGAGAACTTCTCATGATTAAGCTATTGCTTTTGAAG 265
Qy 731 CCATGTCGTATTAAGATGATGATCCCAAGATGAGTGGAGTGGAGTGGAG 790
Db 266 CCATGTCGTATTAAGATGATGATCCCAAGATGAGTGGAGTGGAGTGGAG 325
Qy 791 TTAATGAAAAAGTTCTCAATTTTGAACAAGCTATCAAGATGGCACTATTAATAA 850
Db 326 TTAATGAAAAAGTTCTCAATTTTGAACAAGCTATCAAGATGGCACTATTAATAA 385
Qy 851 ATCTACCTGCTGCTGAGCTGATGAGATTAAGATGATGATGATGATGATGAT 910
Db 386 ATCTACCTGCTGCTGAGCTGATGAGATTAAGATGATGATGATGATGATGAT 445
Qy 911 GATTAGAAAGCCATTCATGCGACAGAGATATTTA 946
Db 446 GATTAGAAAGCCATTCATGCGACAGAGATATTTA 481

RESULT 8
US-09-814-353-3455
; Sequence 3455, Application US/09814353
; Publication No. US20030165831A1

GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3455
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-3455

Query Match 14.0%; Score 434; DB 12; Length 434;
Best Local Similarity 100.0%; Pred.No.4.3e-107;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1186 ACTGAGTGCAGACAGAGAGAGAGAGATCCAGAGTGAAGTGAAGTGAAGTGAAGT 1245
Db 1 ACTGAGTGCAGACAGAGAGAGAGAGATCCAGAGTGAAGTGAAGTGAAGTGAAGT 60
Qy 1246 TTACAGTATTCAGACAGTGAATTTACTGCTGCTTACTGACAGTCTTATAGCTTT 1305
Db 61 TTACAGTATTCAGACAGTGAATTTACTGCTGCTTACTGACAGTCTTATAGCTTT 120
Qy 1306 ACTAGAAAGAGACCAATGCTGTTGAGAGCTCAAAATTAAGTGTTCAGAGCAGAT 1365
Db 121 ACTAGAAAGAGACCAATGCTGTTGAGAGCTCAAAATTAAGTGTTCAGAGCAGAT 180
Qy 1366 CTTCCTTCCTCAATTAATTCATTCATGATGATGATGATGATGATGATGATGAT 1425
Db 181 CTTCCTTCCTCAATTAATTCATTCATGATGATGATGATGATGATGATGATGATGAT 240
Qy 1426 AAAGAGATCATCCATTAATGATGATGATGATGATGATGATGATGATGATGAT 1485
Db 241 AAAGAGATCATCCATTAATGATGATGATGATGATGATGATGATGATGATGAT 300
Qy 1486 CCTACCTTCCCTGATATGCAAAATTAATTAAGGAGAGAAATGTTGATTTTGA 1545
Db 301 CCTACCTTCCCTGATATGCAAAATTAATTAAGGAGAGAAATGTTGATTTTGA 360
Qy 1546 AGATTATTAATGATTAATAAACTGCTGAGCTGTGATTTTAACAAATTAACAT 1605
Db 361 AGATTATTAATGATTAATAAACTGCTGAGCTGTGATTTTAACAAATTAACAT 420
Qy 1606 ATCCTGATTTT 1619
Db 421 ATCCTGATTTT 434

RESULT 9
US-09-814-353-9771
; Sequence 9771, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John

```

; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9771
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-814-353-9771

```

```

Query Match      14.0%; Score 434; DB 12; Length 434;
Best Local Similarity 100.0%; Pred. No. 4.3e-107;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1186 ACTCGAAGTCGACAGAGAGAGAGAGAGATCCAGAAAGTTGAAGTACAGACCAACCAATGT 1245
DB 1 ACTCGAAGTCGACAGAGAGAGAGAGAGATCCAGAAAGTTGAAGTACAGACCAACCAATGT 60
QY 1246 TTGACGATATTCAGCAGAGAGAGAGAGAGATTTACTCGTGTCTTACTGACAGTCTTATAGCCTT 1305
DB 61 TTGACGATATTCAGCAGAGAGAGAGAGAGATTTACTCGTGTCTTACTGACAGTCTTATAGCCTT 120
QY 1306 ACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1365
DB 121 ACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 1366 CTTCCTTTCTGCCATTCATTAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1425
DB 181 CTTCCTTTCTGCCATTCATTAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 240
QY 1426 AAGAGAGATCATCCAAATATGATGGGTTTGAACCCCTTGTGAACGAGAGAGAGAGAGAGAG 1485
DB 241 AAGAGAGATCATCCAAATATGATGGGTTTGAACCCCTTGTGAACGAGAGAGAGAGAGAGAG 300
QY 1486 CTTACCTTCCTCGATATGCAAAAATTAATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1545
DB 301 CTTACCTTCCTCGATATGCAAAAATTAATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 1546 AGATTATATAGATAGATTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1605
DB 361 AGATTATATAGATAGATTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 1606 ATCTGAGATTTT 1619
DB 421 ATCTGAGATTTT 434

```

```

RESULT 10
US-09-983-965-210
; Sequence 210, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.

```

```

; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,211
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 210
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 25-BOWM1-018-Q1-E1-Q1
; US-09-983-965-210

```

```

Query Match      11.3%; Score 350.8; DB 10; Length 398;
Best Local Similarity 94.3%; Pred. No. 1.7e-84;
Matches 364; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

```

QY 2298 GCCGAGATGACATGAGCCAGCATATGAGACATGATGCTGGAATGTTAAACCA 2357
DB 13 GCCGAGATGACATGAGCCAGCATATGAGACATGATGCTGGAATGTTAAACCA 72
QY 2358 TGTAGCATTTGACATGAGCCAGCAAGTACGTAAACCAAGTTGAGCTTGATGTGAAC 2417
DB 73 TGTAGCATTTGACATGAGCCAGCAAGTACGTAAACCAAGTTGAGCTTGATGTGAAC 132
QY 2418 AAGTTGGATGACACACAGGTTTCTCCATTCACACAGTATATACCCCGCCGACAGGC 2477
DB 133 AAGTTGGATGACACACAGGTTTCTCCATTCACACAGTATATACCCCGCCGACAGGC 192
QY 2478 ACTACTACAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2537
DB 193 ACTACTACAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252
QY 2538 CTGAACCTGATGTGACAGCTGTGTAAGCAGCTTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2597
DB 253 CTGAACCTGATGTGACAGCTGTGTAAGCAGCTTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
QY 2598 TTCCCTAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2657
DB 313 TTCCCTAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372
QY 2658 TTATGAGATTATGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2683
DB 373 TTATGAGATTATGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398

```

```

RESULT 11
US-09-783-590-5693
; Sequence 5693, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haselaine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5693

```


	Query Match	8.7%	Score 268.6;	DB 10;	Length 501;
	Best Local Similarity	93.2%	Pred No. 4,5e-62;		
	Matches 313;	Conservative 0;	Mismatches 17;	Indels 6;	Gaps 3;
OY	2223	AAAGCATATGACAGACGCAGAAAGCGCTGTAGTGTAAAAAACAAGA	2283		
Db	7	AAAGCATATGAAAGACGACGAGAAAGCCGTGTGTGTAAAAAACAAGA	66		
OY	2283	AAGTCGCCCAATTGAGCCGAGAGATCACAATGAGCCAAGCATATCAGAACATGTGTGCTG	2344		
Db	67	AAAGTCGCCCAATTGAGCCGAGAGATCACATATAGCCAAGCATATCAGAACATGTGTGCTG	126		
OY	2343	GAATGTTTAAAACCATGCTAGACATTTTGACATGACGGCAAAATACGTAAACCGAAGTTTG	2402		
Db	127	GAATGTTTAAAACCATGCTAGACATTTTGACATGACGGCAAAATACGTAAACCGAAGTTTG	186		

```

Query Match%: 5.9%; Score 182; DB 12; Length 489;
Best Local Similarity: 100.0%; Pred. No. 1.6e-38;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1431 AGATCATCCATTATGATGGGTTTGAACCCCTTGTGAACGAGGCTACTTCCACCTAC 1499
DB      401 AGATCATCCATTATGATGGGTTTGAACCCCTTGTGAACGAGGCTACTTCCACCTAC 342

QY      1491 CTTCCTCGATTCGAAAAATATTAAAAAGGAAAGAAATGCTGAACCTATTTGGCAAGAT 1559
DB      341 CTTCCTCGATTCGAAAAATATTAAAAAGGAAAGAAATGCTGAACCTATTTGGCAAGAT 282

QY      1551 AATAGATGAATPAAAACTGTCTGAGGTTGTAATTACAAATTTACATTGATCT 1619
DB      281 AATAGATGAATPAAAACTGTCTGAGGTTGTAATTACAAATTTACATTGATCT 222

QY      1611 GG 1612
DB      221 GG 220

RESULT 13
US-10-027-633-323193/c
; Sequence 323193, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

```

TITLE OF INVENTION: Polymorphisms in the Human Genome
 FILE REFERENCE: 108827.129
 CURRENT APPLICATION NUMBER: US/10/027,632
 PRIOR FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: US 60/218,006
 PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/198,676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193,483
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: US 60/185,218
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: US 60/167,363
 PRIOR FILING DATE: 1999-11-23
 PRIOR APPLICATION NUMBER: US 60/156,358
 PRIOR FILING DATE: 1999-09-28
 PRIOR APPLICATION NUMBER: US 60/146,002
 PRIOR FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 323193
 LENGTH: 489
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-323193

Query Match
 Best Local Similarity 5.9%; Score 182; DB 13; Length 489;
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1431 ACATCTCCAAATATGATGGTGGTTGCAACCCCTTGGAACCGAGAGCTACTTCCACCTAC 1490
 DB 401 AGATCTATCCAAATATGATGGTGGTTGCAACCCCTTGGAACCGAGAGCTACTTCCACCTAC 342

QY 1491 CTTCCTCGATATGCAAAATAATTAAGGAGAGAAATGCTGAACCTATTTGGCAAGATT 1550
 DB 341 CTTCCTCGATATGCAAAATAATTAAGGAGAGAAATGCTGAACCTATTTGGCAAGATT 282

QY 1551 AATAGTAAATTAACCTGCTGAGAGTTGGAATTTAAACAATTACATTGTATCTT 1610
 DB 281 AATAGTAAATTAACCTGCTGAGAGTTGGAATTTAAACAATTACATTGTATCTT 222

QY 1611 GG 1612
 DB 221 GG 220

RESULT 14
 US-10-001-857-41/c
 Sequence 41, Application US/10001857
 Publication No. US20020183500A1
 GENERAL INFORMATION:
 APPLICANT: Macina, Roberto
 APPLICANT: Recipon, Herve
 APPLICANT: Chen, Sei-Yu
 APPLICANT: Sun, Yongming
 APPLICANT: Liu, Chenghua
 TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro
 FILE REFERENCE: DEX-0273
 CURRENT APPLICATION NUMBER: US/10/001,857
 PRIOR FILING DATE: 2001-11-20
 PRIOR APPLICATION NUMBER: 60/252,054
 PRIOR FILING DATE: 2000-11-20
 NUMBER OF SEQ ID NOS: 208
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 41
 LENGTH: 151
 TYPE: DNA
 ORGANISM: Homo sapien
 US-10-001-857-41

Query Match
 Best Local Similarity 4.5%; Score 140.4; DB 13; Length 151;
 Matches 99.3%; Pred. No. 1.3e-27;

Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1911 CATATCTTGAGGAATTTGCCACCTTGCAAGATGAGTTATGACATTTATTTAATAG 1970
 DB 151 CATATCTTGAGGAATTTGCCACCTTGCAAGATGAGTTATGACATTTATTTAATAG 92

QY 1971 CAGAGAGGTTGATGAGCGCTTCACACCATGCTTTGAACAGAACCCCAAGGCAAC 2030
 DB 91 CAGAGAGGTTGATGAGCGCTTCACACCATGCTTTGAACAGAACCCCAAGGCAAC 32

QY 2031 ATTTGGCTGTTAGTACTTG 2052
 DB 31 ATTTGGCTGTTAGTACTTG 10

RESULT 15
 US-10-001-857-111/c
 Sequence 111, Application US/10001857
 Publication No. US20020183500A1
 GENERAL INFORMATION:
 APPLICANT: Macina, Roberto
 APPLICANT: Recipon, Herve
 APPLICANT: Chen, Sei-Yu
 APPLICANT: Sun, Yongming
 APPLICANT: Liu, Chenghua
 TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P
 FILE REFERENCE: DEX-0273
 CURRENT APPLICATION NUMBER: US/10/001,857
 PRIOR FILING DATE: 2001-11-20
 PRIOR APPLICATION NUMBER: 60/252,054
 PRIOR FILING DATE: 2000-11-20
 NUMBER OF SEQ ID NOS: 208
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 111
 LENGTH: 815
 TYPE: DNA
 ORGANISM: Homo sapien
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (206)..(305)
 OTHER INFORMATION: a, c, g or t
 US-10-001-857-111

Query Match
 Best Local Similarity 2.8%; Score 87.8; DB 13; Length 815;
 Matches 122; Conservative 0; Mismatches 2; Indels 23; Gaps 1;

QY 1910 TCATATCTTGAGGAATTTGCCACCTTGCAAGATGAGTTATGACATTTATTTAATAG 1969
 DB 205 TCATATCTTGAGGAATTTGCCACCTTGCAAGATGAGTTATGACATTTATTTAATAG 169

QY 1970 GCAGAGAGGTTGATGAGCGCTTCACACCATGCTTTGAACAGAACCCCAAGGCAAC 2029
 DB 168 GCAGAGAGGTTGATGAGCGCTTCACACCATGCTTTGAACAGAACCCCAAGGCAAC 109

QY 2030 CATTGGCCTGTTAGTACTTG 2056
 DB 108 CATTGGCCTGTTAGTACTTG 82

Search completed: November 25, 2003, 03:08:51
 Job time: 938 secs

us-10-001-857-42.rng

Page 1

```

OM nucleic - nucleic search, using sw model
Run on:      November 24, 2003, 21:45:44
              (without alignments)
              10728.447 Million cell updates/sec

```

```

Title: US-10-001-857-42
Perfect score: 3096
Sequence: 1 ttcttcacgaactccag.....attcttgacaaaaaaaa 3096

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%

Listing first 45 summaries

```
Database :
_N_Genseq 13Jun03.*
1: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
8: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
9: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
10: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
11: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
12: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
13: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
14: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
15: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
16: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
17: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
18: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
19: /SIDS1/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
20: /SIDS1/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
21: /SIDS1/gcgdata/genseq/genseqn-emb1/NA2001.DAT.*
22: /SIDS1/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*
23: /SIDS1/gcgdata/genseq/genseqn-emb1/NA2003.DAT.*
24: /SIDS1/gcgdata/genseq/genseqn-emb1/NA2004.DAT.*
25: /SIDS1/gcgdata/genseq/genseqn-emb1/NA2005.DAT.*
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Query No.	Score	Match	Length	DB	ID	Description
1	3096	100.0	3096	24	AB075303	Human lung specific
2	2577.8	83.3	2668	23	AB523940	Human prostate expt
3	2577.8	83.3	2668	23	ABV29823	Human prostate expt
4	2417.4	78.1	2488	24	AAI44929	Human molecule for
5	717.4	23.2	719	24	AB551309	cDNA encoding human
6	30.8	11.3	398	25	AB502081	Bovine EST associa
7	343.8	11.1	799	24	AB051658	Oligonucleotide f
8	343.8	11.1	799	24	AB051659	Oligonucleotide f

C	9	293.1	9.4	356	23	ABV15531	Human prostate exp
C	10	288.8	9.3	439	23	ABV326129	Human prostate exp
C	11	288.8	9.3	439	23	ABV45182	Human prostate exp
C	12	274.2	8.9	799	24	ABO51160	Oligonucleotide fo
C	13	274.2	8.9	799	24	ABO51161	Oligonucleotide fo
C	14	230.2	7.4	303	23	ABV06162	Human prostate exp
C	15	217	7.0	2515	23	ABV07143	Drosophila melanog
C	16	140.4	4.5	151	24	ABO75302	Human lung specifi
C	17	90.6	2.9	3401	23	ABV05906	Drosophila melanog
C	18	90.6	2.9	4847	23	ABV07142	Human lung specifi
C	19	87.8	2.8	815	24	ABO75537	Human spliced tran
C	20	60	1.9	600	24	ABN43747	Oligonucleotide fo
C	21	58.6	1.9	600	24	ABO52446	Oligonucleotide fo
C	22	58.6	1.9	600	24	ABO52447	Oligonucleotide fo
C	23	56.4	1.8	320	21	AAA81813	Primer used in the
C	24	56	1.8	451	24	ABT10409	Human breast cance
C	25	56	1.8	1286	23	ABV19656	Mouse ischaemic co
C	26	55.4	1.8	365	24	ABV54856	Human prostate exp
C	27	55.2	1.8	320	21	AAA81816	Primer used in the
C	28	54.8	1.8	320	21	AAA81815	Primer used in the
C	29	54.6	1.8	840	24	ABO35444	Oligonucleotide fo
C	30	54.6	1.8	840	24	ABO35445	Oligonucleotide fo
C	31	54.6	1.8	1416	25	ABZ20967	Animal test kit ma
C	32	54.6	1.8	12733	24	ABK98651	Vector pEGFP1 con
C	33	54.6	1.8	13739	24	ABK98652	Vector pEGFP1 con
C	34	54.2	1.8	318	21	AAA81814	Primer used in the
C	35	54.2	1.8	4403765	22	AAV19663	Mycobacterium tub
C	36	54.2	1.8	4411529	22	AAV19668	Mycobacterium tub
C	37	53.8	1.7	2188	20	AAZ775062	Human ovarian tumo
C	38	53.4	1.7	1337	20	AAZ17263	Human gene express
C	39	53.4	1.7	2561	22	AAH26500	Rabbit low density
C	40	53	1.7	434	22	AAH44343	Human polynucleoti
C	41	52.8	1.7	344	22	AAH82007	Human polynucleoti
C	42	52.2	1.7	1327	24	ABO68452	Listeria monocytog
C	43	51.6	1.7	1000	21	AAA02484	Human colon cancer
C	44	51.4	1.7	237	21	AAA81807	N. meningitidis pa
C	45	51	1.6	712	24	ABO40856	Oligonucleotide fo

ALIGNMENTS

	RESULT
ABQ75303	ID
ABQ75303 standard; cDNA; 3096 BP.	
AC	ABQ75303;
DT	05-NOV-2002 (first entry)
DE	Human lung specific nucleic acid sequence SEQ ID NO:42.
KX	Human; lung; lung specific nucleic acid; LSN; lung specific protein
KW	LSP; cytosolic; gene therapy; vaccine; metastasis; lung cancer;
OS	squamous cell carcinoma; gene; chromosome 9; ss.
PN	Homo sapiens.
WO200264788-A2.	
22-AUG-2002.	
20-NOV-2001; 2001WO-US45080.	
20-NOV-2000; 2000US-252054P.	
(DIAD-) DIADEXUS INC.	
Maccina RA, Recipon H, Chen S, Sun Y, Liu C;	
WPI; 2002-657601/70.	
New lung specific nucleic acid useful in gene therapy or as vaccines	

661 TTGAAGAGCTTGTCAGATTAAAGTGGAGACACTTCATGATAAGCTATTGGT 720
 721 CTTTGAAGCATGCTGCTATTGAATGATGATCCCAAGATGATGCTGSCATATT 780
 781 CTTTTCCTTCCATCTCTCTATTGAATGATGATCCCAAGATGATGCTGSCATATT 780

Db	1/2	CTTTTGGAGGCGCATTTG	84
Qy	781	GGAAACCAAGTATCGAAAGTTCTCAATTTTGAACAAGGATCAAGATGAGCAGTATT	84
Db	781	GGAAACCAAGTATCGAAAGTTCTCAATTTTGAACAAGGATATAGAGTGGCAGTATT	84
Qy	841	AAAAATTTAAATATCTCACTGCTGAGCTGAGTATGAGATATGATATCATGTTTTTGCTGT	90
Db	841	AAAAATTTAAATATCTCACTGCTGAGCTGAGTATGAGATATGATATCATGTTTTTGCTGT	90
Qy	901	TTGATTAACGCTGTAGAAAGCCATTCATCTGGCAGCAGACGATTTTACGCGCTTTACATT	96
Db	901	TTGATTAACGCTGTAGAAAGCCATTCATCTGGCAGCAGACGATTTTACGCGCTTTACATT	96
Qy	961	CATAATCCAGACTTTATAGAAAGATCCGCGATATGAAAGCTTTTGCTCTGGAAATCTTGAA	10
Db	961	CATAATCCAGACTTTATAGAAAGATCCGCGATATGAAAGCTTTTGCTCTGGAAATCTTGAA	10

1021 ATCTGTGACATTGCATAGGGAAAAAGTAAATAAAGCTGTGTTTGGAAAGAGGAGGAAA... 2

Db 1021 ATCTGTGACATTGCGAAGGAAAAAGTAAATTAACCTGCTGTTTTCGACAGAGGAGGATTA...

Qv 1081 CAGTCAATGACTTATGCAATTTAAATGGCTAACAGTGTGACAGATCTTCGAGTTACAGGC 11

Db	1081	CAGTCATGACTTATGGATTAAATGGCTAACGCTGAGAGTCTTCGAGTACAGGC	1140
Qy	1141	ATGCTAAABAATGCGGAGATGACATGCCAAAGAGTAAAGATCTCGAAGTCACAA	1200
Db	1141	ATGCTAAAGATGCGAGATGACATGCCAAAGAGATAAAGATCTCGAAGTCACAA	1200
Qy	1201	GGAGAGAAAGAGATCCAGAGTTGAACATAGAACCCAAACAATGTTAGCAGTATTCAGC	1260
Db	1201	GGAGAGAAAGAGATCCAGAGTTGMACTAACAACCAACATGTTTAGCAGTATTCAGC	1260
Qy	1261	AGAGTGAATTTACTCCGTGTGTACTGCAGTGCCTTATAGCCTTTACTAAGAAAGAC	1320
Db	1261	AGAGGAAATTTACTCCGTGTGTACTGCAGTGCCTTATAGCCTTTACTAAGAAAGAC	1320
Qy	1321	AGTGTGTGGAGAGCTCAAAAATTGATGTTCAAGCAGCAGATCTTCTTGCCATT	1380
Db	1321	AGTGTGTGGAGAGCTCAAAAATTGATGTTCAAGCAGCAGATCTTCTTGCCATT	1380
Qy	1381	CATATTTCAATGCATCATGGCATCCAGGCCCAAGATGATACTACAAAGAGATCATCCA	1440
Db	1381	CATATTTCAATGCATCATGGCATCCAGGCCCAAGATGATACTACAAAGAGATCATCCA	1440
Qy	1441	ATTATGATGGTTTTGGAACTCCCTGTGAAACCAAGGCTACTTCGACCTACCTCCCTCGA	1500
Db	1441	ATTATGATGGTTTTGGAACTCCCTGTGAAACCAAGGCTACTTCGACCTACCTCCCTCGA	1500
Qy	1501	TATGCAAAATATTTAAAGGGAAGAAATGATGATATTTTGCAGATTATATGATAGA	1560
Db	1501	TATGCAAAATATTTAAAGGGAAGAAATGATGATATTTTGCAGATTATATGATAGA	1560
Qy	1561	ATPAAAACGTCTGTGAGGTTTGAAATTTAACAAATTTAATGTATCCGAGATTTTTC	1620
Db	1561	ATPAAAACGTCTGTGAGGTTTGAAATTTAACAAATTTAATGTATCCGAGATTTTTC	1620
Qy	1621	TGTGAATTTAGTGAACAGTCACCATGTGTTCTTTCAAGATCTCTGTTCAAAACCATCTTC	1680
Db	1621	TGTGAATTTAGTGAACAGTCACCATGTGTTCTTTCAAGATCTCTGTTCAAAACCATCTTC	1680
Qy	1681	CTGGTGATACAAAAGGTCTTTGGAACTCATCTCATGCAGACATGTGTGAAGATGCA	1740
Db	1681	CTGGTGATACAAAAGGTCTTTGGAACTCATCTCATGCAGACATGTGTGAAGATGCA	1740
Qy	1741	CTTGCGCTTTTGCATCCTCCGAGTGTCTTCCCCCAAGTGTACCTATATATATC	1800

Db 1741 CTTGGCTCTTTGTAGATCTCCGAGTCTTCCCGCAAGTCACTATATATATATC 1800
 Qy 1801 ACCAGGCTAAGAGCTGTATGACATCTCTTGTACTACTGTGTGGCCATTCTGTAGTC 1860
 Db 1801 ACCAGGCTAAGAGCTGTATGACATCTCTTGTACTACTGTGTGGCCATTCTGTAGTC 1860
 Qy 1861 TATATGAGATCCATGAGACATTAACAGGCTCGACAGAGATTAAGCTTGTCTATTTCTTG 1920
 Db 1861 TATATGAGATCCATGAGACATTAACAGGCTCGACAGAGATTAAGCTTGTCTATTTCTTG 1920
 Qy 1921 AGGAATTTGCCACCTTGACAGATGAGTTATGACATTTATTTATTAAGCAGAGAGGT 1980
 Db 1921 AGGAATTTGCCACCTTGACAGATGAGTTATGACATTTATTTATTAAGCAGAGAGGT 1980
 Qy 1981 TGATGAGGCGCTTACACCATGCTGTGAAACAGAACCCCAAGGCACTTTGGCTG 2040
 Db 1981 TGATGAGGCGCTTACACCATGCTGTGAAACAGAACCCCAAGGCACTTTGGCTG 2040
 Qy 2041 TTTAGTACCTGGTCTCTTACATTAACCTTGCATTTATGATACAGTACCTTGAAGTG 2100
 Db 2041 TTTAGTACCTGGTCTCTTACATTAACCTTGCATTTATGATACAGTACCTTGAAGTG 2100
 Qy 2101 CTTTGAATTTGAACTTACAGTATGACAGATCTATTAATATTTGATTTCTCTGA 2160
 Db 2101 CTTTGAATTTGAACTTACAGTATGACAGATCTATTAATATTTGATTTCTCTGA 2160
 Qy 2161 ATTCCTTTACGATGTTGATGTCAACATTTGAGTGTGCCATGGCTCTCAATAGGCGA 2220
 Db 2161 ATTCCTTTACGATGTTGATGTCAACATTTGAGTGTGCCATGGCTCTCAATAGGCGA 2220
 Qy 2221 GGAAGGATTAAG 2280
 Db 2221 GGAAGGATTAAG 2280
 Qy 2281 GAAAGTTGCGCCATTGAGCGAGAGATCAATGAGCCAGATTCAGAACTATGTGTC 2340
 Db 2281 GAAAGTTGCGCCATTGAGCGAGAGATCAATGAGCCAGATTCAGAACTATGTGTC 2340
 Qy 2341 TGGATGTTTAAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
 Db 2341 TGGATGTTTAAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
 Qy 2401 TGAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460
 Db 2401 TGAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460
 Qy 2461 GACCCCGCGCGAGTGCATCTTACAGTTCAGAGAAATGTCTGACCTCAATTAATATAG 2520
 Db 2461 GACCCCGCGCGAGTGCATCTTACAGTTCAGAGAAATGTCTGACCTCAATTAATATAG 2520
 Qy 2521 CCCCTCTCTCAGTCTCTGAACTGTATGTGAGAGTGTAGCACTTTCAACAGCGAAA 2580
 Db 2521 CCCCTCTCTCAGTCTCTGAACTGTATGTGAGAGTGTAGCACTTTCAACAGCGAAA 2580
 Qy 2581 AATGATATGAAAAATATCTCTTAAACCCGAGCATGAGTTAATGAATTTTAAAGGTTC 2640
 Db 2581 AATGATATGAAAAATATCTCTTAAACCCGAGCATGAGTTAATGAATTTTAAAGGTTC 2640
 Qy 2641 CAACCCCACTTTGTGTTATGAGTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
 Db 2641 CAACCCCACTTTGTGTTATGAGTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
 Qy 2701 TCCCTCTGATTTGATTTCTCTGCTCAATTAATTTTCTGTTGTGAACCTTTGAGAG 2760
 Db 2701 TCCCTCTGATTTGATTTCTCTGCTCAATTAATTTTCTGTTGTGAACCTTTGAGAG 2760
 Qy 2761 GAGACTGGAGAGTGGCCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
 Db 2761 GAGACTGGAGAGTGGCCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
 Qy 2821 CACATCACCAGGCTCACAATCAGGAAAGTGAATTTCTTGGTTAACACCTCANTA 2880

Db 2821 CACATCACCAGGCTCACAATCAGGAAAGTGAATTTCTTGGTTAACACCTCANTA 2880
 Qy 2881 TAAGAAATACCTTTAGTTTGGACAGCTTATATGACATGATGAAGAACTGCTGTTTAAAG 2940
 Db 2881 TAAGAAATACCTTTAGTTTGGACAGCTTATATGACATGATGAAGAACTGCTGTTTAAAG 2940
 Qy 2941 TGGTTTATATGTTTCATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
 Db 2941 TGGTTTATATGTTTCATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
 Qy 3001 GGTTTTATATGATTTATATCAAAATCATTTTATGATGATGATGATGATGATGATGATG 3060
 Db 3001 GGTTTTATATGATTTTATATCAAAATCATTTTATGATGATGATGATGATGATGATGATG 3060
 Qy 3061 TTAATAAGGTTAATAATTTCTTGAACAAAAA 3096
 Db 3061 TTAATAAGGTTAATAATTTCTTGAACAAAAA 3096
 Db 3061 TTAATAAGGTTAATAATTTCTTGAACAAAAA 3096
 RESULT 2
 ABV23940/c
 ID ABV23940 standard; cDNA; 2668 BP.
 XX
 AC ABV23940;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 23931.
 XX
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens
 PN W0200160860-A2.
 XX
 PD 23-APR-2001
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-21314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX
 PS Claim 1; Page 4423; 11750bp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (1) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (1) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient.

Tue Nov 25 09:35:51 2003

us-10-001-857-42.rng

Page 4

(T) is also useful as a pharmacodynamic or pharmacogenomic marker.

sequence 2668-BP: 742 A; 617 C; 483 G; 820 T; 6 other

83.3%; Score 2577.8; DB 23; Length 2668

```

Query match...      Pred. No. 0;
Best Local Similarity 98.8%;
Matches 2638; Conservative 0; Mismatches 7; Indels 26; Gaps 3

```

[illegible]

Db	1708	TTCAATTCATCATGCAATCCAGGCCCAAGATGATACCTAATAAGAGATCATCCAAATTAT	1649
Qy	1446	GATAGGTTTAAACCCCTTTGTGAACCAAGGCTACTCCACTACCTCCCTCGATATGC	1505
Db	1648	GATGGTTTGAACCCCTTTGTGAACCAAGGCTACTCCACTACCTCCCTCGATATGC	1589
Qy	1506	AAAAATTAATTAAGGGAAGAAATGGTGAACATATTTGCAGATTATAGATGAATAAA	1565
Db	1588	AAAAATTAATTAAGGGAAGAAATGGTGAACATATTTGCAGATTATAGATGAATAAA	1529
Qy	1566	AACGTCTGTGAGTTGTGAATTTAAACAAATTTACATTTGATACCTGATTTTTCTGTGA	1625
Db	1528	AACGTCTGTGAGTTGTGAATTTAAACAAATTTACATTTGATACCTGATTTTTCTGTGA	1469
Qy	1626	ATTATGTGAACAGTCAACCATGTGTTCTTCAAGATCTCTGTGTAAACCACTTTCCTG	1685
Db	1468	ATTATGTGAACAGTCAACCATGTGTTCTTCAAGATCTCTGTGTAAACCACTTTCCTG	1409
Qy	1686	GGATAACAAAAGGTCCTTTGGAACTCATCTCATGCAAGACATGTGTAAATATGACACTCG	1745
Db	1408	GGATAACAAAAGGTCCTTTGGAACTCATCTCATGCAAGACATGTGTAAATATGACACTCG	1349
Qy	1746	GTCCTTTGTGCAGATCCTCCGAGTGGCTTTCCCCCAAGTGTACCTATATAATATCACGAG	1805
Db	1348	GTCCTTTGTGCAG--TCTCCGGTGTCTTTCCCCCAAGTGTACCTATATAATATCACGAG	1291
Qy	1806	GCTAAGACTGTATCGACTCTCTTTGTATCTCATCTGTGTGGCCATCTGTATAGCTTAT	1865
Db	1290	GCTAAGACTGTATCGACTCTCTTTGTATCTCATCTGTGTGGCCATCTGTATAGCTTAT	1231
Qy	1866	CAGATCCATGACATTAACAGGGCTCGACAGAGATTAAGCTTGTGATATTTCTTGAGGAA	1925
Db	1230	CAGATCCATGACATTAACAGGGCTCGACAGAGATTAAGCTTGTGATATTTCTTGAGGAA	1171
Qy	1926	TTTACCACTTGCAGAGATGATTATGACATTTTATTTAAATAGCCAAAGAAAGTTGATG	1985
Db	1170	TTTGCACCTTGCAGAGATG-----AGCCAAAGAAAGTTGATG	1134
Qy	1986	CAGGCTTCAACACATGCTGTTGAAACAGAAACCCCAAGGCACATTTGGCTGTTTATG	2045
Db	1133	CAGGCTTCAACACATGCTGTTGAAACAGAAACCCCAAGGCACATTTGGCTGTTTATG	1074
Qy	2046	GTACCTGGGCTCTTACCATTAACCTTCCGATTATGATACAGTACCTTCAAGTGGCTTGG	2105
Db	1073	GTACCTGGGCTCTTACCATTAACCTTCCGATTATGATACAGTACCTTCAAGTGGCTTGG	1014
Qy	2106	AATTTGAACCTACAGATATGACAGAG--TACTATTAACATATATTTGGTATCTCTGAATTC	2164
Db	1013	AATTTGAACCTACAGATATGACAGAGATCTATTAATATATTTGGTATCTCTCTGAATTC	954
Qy	2165	CTTTACGATGTTGATGTCAACATTGTAAGTCCGCGCGATGGGCTCTCAAAATGGCAGAGAA	2224
Db	953	CTTTACGATGTTGATGTCAACATTGTAAGTCCGCGCGATGGGCTCTCAAAATGGCAGAGAA	894
Qy	2225	AGGATATATGGAAGACAGCAGAAAGGCCGTAGTATGTAATAAAAAAATAAGAAAAA	2284
Db	893	AGGATATATGGAAGACAGCAGAAAGGCCGTAGTATGTAATAAAAAAATAAGAAAAA	834
Qy	2285	GTTTCGCCCATTTAGCCGAGAGATCACAATGAGCCAAGCATATGCAACATGTGTGCTGA	2344
Db	833	GTTTCGCCCATTTAGCCGAGAGATCACAATGAGCCAAGCATATGCAACATGTGTGCTGA	774
Qy	2345	ATGTTTAAAAACAATGTAGCATTTGACATGACGCGAAAGTACGTAAACCGAATTTGAG	2404
Db	773	ATGTTTAAAAACAATGTAGCATTTGACATGACGCGAAAGTACGTAAACCGAATTTGAG	714
Qy	2405	CTTGATATGTGAACAAGTTCCGTATGAAACAAGGTTTGCTCATTTCAACAGGTATGATGC	2464
Db	713	CTTGATATGTGAACAAGTTCCGTATGAAACAAGGTTTGCTCATTTCAACAGGTATGATGC	654
Qy	2465	CCGCGCCAGTGCACACTTACCAAGTTCAAGAAAAATGTTGACCTTCAATTAATATAGCCCT	2522

Db		653	CCGCCGCGACGTGACTTACTTACAGTTTCMAGAAAATGCTGCACCTCAATAAATATAGCCT	594
Oy		2525	CCTCCTCAGCTCCTCTGAAGTGTATGTGGCAGCTTAGTAAGCACTTTCAACAGCGAAAAAG	2584
Db		593	CTCCTCAGTCTCCTGAACTGTATGTGGCAGCTTAGTAAGCACATTTTAAACGGCAAAAATG	534
Oy		2585	AATATTGGAAAAATATCTCTTAACC GGACCATGAGGTTAATGAATTTTAAAGGTGGCAA	2644
Db		533	ATATTGGAAAAATATCTCTTAACC GGACCATGAGGTTAATGAATTTTAAAGGTGGCAA	474
Oy		2645	CCCAACTTGCGTTATGAAGTTATTTGGCAGAGACACA AAAAGGAATCTAAAGTTCT	2704
Db		473	CCCAACTTGCGTTATGAAGTTATTTGGCAGAGACACA AAAAGGAATCTAAAGTTCT	414
Oy		2705	CCTGAATTTGATTTCTCTGCTCATTAATATTTTCTGTGGAAAATTGTGTAGAGAGA	2764
Db		413	CCTGAATTTGATTTCTCTGCTCATTAATATTTTCTGTGGAAAATTGTGTAGAGAGA	354
Oy		2765	CTGGGAGAGTGGCCATAAGAGGGCAGAGCTTCTTTCAGACCACCTTTAGAGGGACA	2824
Db		353	CTGGGAGAGTGGCCATAAGAGGGCAGAGCTTCTTTCAGACCACCTTTAGAGGGACA	294
Oy		2825	TCACCAGGCTCCACATCACGGGAAGTGAGATTTCTTGGGTAACACCTCATTAAG	2884
Db		293	TCACCAGGCTCCACATCACGGGAAGTGAGATTTCTTGGGTAACACCTCATTAAG	234
Oy		2885	GAACTCTTTAGTTTGACAGCCTTATATGACATGATGAAAACTGCTTTTAAAGTGT	2944
Db		233	GAACTCTTTAGTTTGACAGCCTTATATGACATGATGAAAACTGCTTTTAAAGTGT	174
Oy		2945	TTATTATGTTCCATGAGAAAGAACTGCTTATTATGATGATGAAAGTTATATGTT	3004
Db		173	TTATTATGTTCCATGAGAAAGAACTGCTTATTATGATGATGAAAGTTATATGTT	114
Oy		3005	TTATTACAGATTAAATCACAAATCATTTTATGATGATGAGTGAATGATGTTAT	3064
Db		113	TTATTACAGATTAAATCACAAATCATTTTATGATGATGAGTGAATGATGTTAT	54
Oy		3065	AAAGGTTAATAAATTTCTTGACAAAAAAA 3095	
Db		53	AAAGGTTAATAAATTTCTTGACAAAAAAA 23	
 RESULT 3 ABV29823/c				
ID	ABV29823	standard;	CDNA; 2668 BP.	
XX	XX	ABV29823;		
AC	16-SEP-2002	(first entry)		
DT	Human prostate expression marker	CDNA 29814.		
DE	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;			
KW	pharmacogenomic marker; gene; ss.			
OS	Homo sapiens.			
PX	WO200160860-A2*			
PD	23-AUG-2001.			
PF	20-FEB-2001; 2001MO-US05171.			
PR	17-FEB-2000; 2000US-183319P.			
PR	16-MAR-2000; 2000US-189862P.			
PR	25-MAY-2000; 2000US-207454P.			
PR	09-JUN-2000; 2000US-211314P.			
PR	18-JUL-2000; 2000US-219007P.			
PR	13-DEC-2000; 2000US-255281P.			
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.			
XX	XX			

Query	Subject	Score	Length	Mismatches	Indels	Gaps
1	Schlegel R, Endege WO, Monahan JE;	83.3%	2577.8	DB 23	Length 2668	
2	WPI; 2001-662795/76.	98.8%	Pred. No. 0			
3	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer	0	Mismatches 7	Indels 26	Gaps 3	
4	Claim 1; Page 6420; 11750pp; English.					
5	The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:					
6	(a) assessing whether a patient is afflicted with prostate cancer;					
7	(b) monitoring the progression of prostate cancer in a patient;					
8	(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;					
9	(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;					
10	(e) selecting a composition for inhibiting prostate cancer in a patient;					
11	(f) assessing the prostate cell carcinogenic potential of a compound;					
12	(g) determining whether prostate cancer has metastasized in a patient;					
13	(h) assessing the aggressiveness or indolence of prostate cancer in a patient;					
14	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.					
15	Sequence 2668 BP; 742 A; 617 C; 483 G; 820 T; 6 other;					
16	Query Match	83.3%	Score 2577.8	DB 23	Length 2668	
17	Best Local Similarity	98.8%	Pred. No. 0			
18	Best Match 2638; Conservative	0	Mismatches	7	Indels	26
19						Gaps 3
20	426 CGCATGCGTSCACGCTCCCGGTGCGGCTGAGCGGAGGAGGCGGCGCGCCGA	485				
21	2668 CGCATGCGTSCACGCTCCCGGTGCGGCTGAGCGGAGGAGGCGGCGCGCCGA	2609				
22	486 GCGCGCGTTCGTTATTTCCGTGCTCGGACAGTGGTGGCGGCGGATGACCA	545				
23	2608 GCGCGCGTTCGTTATTTCCGTGCTCGGACAGTGGTGGCGGCGGATGACCA	2549				
24	546 AGTAGGCGATTAATGTTATGAAAGCTTCTGATGATGATGATGATGATGATGATG	605				
25	2548 AGTAGGCGATTAATGTTATGAAAGCTTCTGATGATGATGATGATGATGATGATG	2489				
26	606 TATGCCAGAAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	655				
27	2488 TATGCCAGAAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2429				
28	666 AGAAGCTTGTGAGAGATTTAAAGTTGGGAGAACTACTTCTATGATTAAGCTATT	725				
29	2428 AGAAGCTTGTGAGAGATTTAAAGTTGGGAGAACTACTTCTATGATTAAGCTATT	2369				
30	726 TGAAGCCATGTTGCTATTTGAAATGATGATGATGATGATGATGATGATGATGATG	785				
31	2368 TGAAGCCATGTTGCTATTTGAAATGATGATGATGATGATGATGATGATGATGATG	2309				
32	786 CGAAGTTATCGMAAAGTTCTCAATTTTGAACAAAGTATCAAGATGACATTAATAAT	845				
33	2308 CGAAGTTATCGMAAAGTTCTCAATTTTGAACAAAGTATCAAGATGACATTAATAAT	2249				
34	846 TAAAGATCTCAGCTTCCCTGAGACTGATGAGGATTAAGGATTAAGGATTAAGGATTAAGG	905				
35	2248 TAAAGATCTCAGCTTCCCTGAGACTGATGAGGATTAAGGATTAAGGATTAAGGATTAAGG	2189				
36	906 AACGCTGTAGAAAGGCGATTCATGCGACAGACAGATATTACGTCCTTTACATTCATAA	965				
37	2188 AACGCTGTAGAAAGGCGATTCATGCGACAGACAGATATTACGTCCTTTACATTCATAA	2129				
38	966 TCCAGACTTATAGAAAGATCTGCTATGAAAGCTTTTGGCTCTGGGATCTTGAATACTG	1025				
39	2128 TCCAGACTTATAGAAAGATCTGCTATGAAAGCTTTTGGCTCTGGGATCTTGAATACTG	2069				
40	1026 TGACATTGCAAGGAGAAAAATAAAGTGTGCTGTTTGAAGAGGAGATTTTCAGTC	1085				

Db 1261 AGATGACCTTGGCTCTTTTGTCACTCTCCG--GTGCTTCCCCCAAGTGTACTATAT 1318
 QY 1794 AATATACACAGGCTAAGAGCTGATGACTCTTTTGTACTGCTGTGTGGCCATTC 1853
 Db 1319 AATATACACAGGCTAAGAGCTGATGACTCTTTTGTACTGCTGTGTGGCCATTC 1378
 QY 1854 TGTAGCTTTATTCAGATCCATGACATACAGGCTCCACAGAGAGATTAAGCTGTTCAT 1913
 Db 1379 TGTAGCTTTATTCAGATCCATGACATACAGGCTCCACAGAGAGATTAAGCTGTTCAT 1438
 QY 1914 ATTCTTGAGGAATTTGCCACCTTGGAGATGAGTTATGACATTTTATTTAATAGGAG 1973
 Db 1439 ATTCTTGAGGAATTTGCCACCTTGGAGATGAGTTATGACATTTTATTTAATAGGAG 1475
 QY 1974 AGAAGGTTATGAGAGGCTTTCACACATGCTGTGAAAAGAGAACCCCAAGGCAACAT 2033
 Db 1476 AGAAGGTTATGAGAGGCTTTCACACATGCTGTGAAAAGAGAACCCCAAGGCAACAT 1535
 QY 2034 TGGCCTGTTAGTACCTGGGCTCTTACCATTAACCTTGGCATTAATGATACAGTACCTTC 2093
 Db 1536 TGGCCTGTTAGTACCTGGGCTCTTACCATTAACCTTGGCATTAATGATACAGTACCTTC 1595
 QY 2094 TAACTGCTTTGAATGGAATCTACAGTATGACAGGATCTATTACATATATTGGTATC 2153
 Db 1596 TAACTGCTTTGAATGGAATCTACAGTATGACAGGATCTATTACATATATTGGTATC 1655
 QY 2154 TCTCTGAATTCCTTACGATGTTGATGTCACATTAAGTGGTGGTGGCTCTTCAAA 2213
 Db 1656 TCTCTGAATTCCTTACGATGTTGATGTCACATTAAGTGGTGGTGGCTCTTCAAA 1715
 QY 2214 TGGCAGAGGAAGATTAATGGAAGAGCAGAGAAAGCCGAGTAAATTAATTAATTA 2273
 Db 1716 TGGCAGAGGAAGATTAATGGAAGAGCAGAGAAAGCCGAGTAAATTAATTAATTA 1775
 QY 2274 AAAAAAGAAAGTTGGCCATTTGAGCCGAGAGATCACAATGAGCCACATATGAGAA 2333
 Db 1776 AAAAAAGAAAGTTGGCCATTTGAGCCGAGAGATCACAATGAGCCACATATGAGAA 1835
 QY 2334 TGTGTCTGGAATGTTTAAACCATGTTGATGATGATGATGATGATGATGATGATGAT 2393
 Db 1836 TGTGTCTGGAATGTTTAAACCATGTTGATGATGATGATGATGATGATGATGATGAT 1895
 QY 2394 CGAAGTTGAGCTGTTAGTGAACAAGTGGTATGACACAGGTTTGTCTCATTCACA 2453
 Db 1896 CGAAGTTGAGCTGTTAGTGAACAAGTGGTATGACACAGGTTTGTCTCATTCACA 1955
 QY 2454 GTGTGATGACCCCGCCGCGCAGTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2513
 Db 1956 GTGTGATGACCCCGCCGCGCAGTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2015
 QY 2514 AATATAGCT 2573
 Db 2016 AATATAGCT 2075
 QY 2574 AGGCAAAATGATATTGAAAATATCTCTTAAACCCGAGCAATGAGTTTAAATTTTAA 2633
 Db 2076 AGGCAAAATGATATTGAAAATATCTCTTAAACCCGAGCAATGAGTTTAAATTTTAA 2135
 QY 2634 AGGTTGCCAAACCCCACTTGTGTGATGATGATGATGATGATGATGATGATGATGAT 2693
 Db 2136 AGGTTGCCAAACCCCACTTGTGTGATGATGATGATGATGATGATGATGATGATGAT 2195
 QY 2694 CTAAAGTCTCTCTGATTTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2753
 Db 2196 CTAAAGTCTCTCTGATTTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2255
 QY 2754 TTTGAGAGACTGGGAGGTTGGCCATTAAGGGGAGAGTCTTTCTTACAGCCCACTCT 2813
 Db 2256 TTTGAGAGACTGGGAGGTTGGCCATTAAGGGGAGAGTCTTTCTTACAGCCCACTCT 2315
 QY 2814 TAGAGGACATACAGGCTCCACATCGGGAAGTGAATGATTTCTTGGGTAAACA 2873
 Db 2316 TAGAGGACATACAGGCTCCACATCGGGAAGTGAATGATTTCTTGGGTAAACA 2375

QY 2874 CTCATTATAGAAATCTTTTATGTTGACAGCTTATATGACATGAATGAATGCTGT 2933
 Db 2376 CTCATTATAGAAATCTTTTATGTTGACAGCTTATATGACATGAATGAATGCTGT 2435
 QY 2934 TTTAAAGTGTATTATGTTCCATGAGAAAGTGTCTTATGAT 2981
 Db 2436 TTTAAAGTGTATTATGTTCCATGAGAAAGTGTCTTATGAT 2483
 RESULT 5
 ABS51309
 ID ABS51309 standard; CDNA: 719 BP.
 XX
 AC ABS51309;
 DT 21-OCT-2002 (first entry)
 XX
 DE
 XX
 CDNA encoding human secretory protein #7.
 Human; secretory polypeptide; SPTM; actinic keratosis; arteriosclerosis;
 bursitis; cirrhosis; hepatitis; polycythaemia vera; anaemia; psoriasis;
 primary thrombocytopenia; cancer; adenocarcinoma; leukaemia; myeloma;
 primary immune system disorder; acquired immunodeficiency syndrome;
 sarcoma; asthma; Crohn's disease; diabetes mellitus; gout;
 AIDS; allergy; asthma; Goodpasture's syndrome; thyroiditis; pancreatitis;
 glomerulonephritis; osteoporosis; Reiter's syndrome;
 hepatitis; multiple sclerosis; neurological disorder; epilepsy; stroke; dementia;
 rheumatoid arthritis; neurologic disease; Huntington's disease; mood; anxiety;
 Alzheimer's disease; Pick's disease; Huntington's disease; mental disorder;
 Parkinson's disease; central nervous system disorder; transgenic animal;
 schizophrenia disorder; amnesia; Tourette's disorder;
 gene therapy; gene; ss.
 KW
 KM
 OS Homo sapiens.
 XX
 XX WO200257304-A2.
 EN 25-JUL-2002.
 PD
 XX
 XX 15-JAN-2002; 2002MO-US01340.
 PF
 XX
 XX 16-JAN-2001; 2001US-261864P.
 PR 16-JAN-2001; 2001US-261865P.
 PR 16-JAN-2001; 2001US-261979P.
 PR 16-JAN-2001; 2001US-261981P.
 PR 17-JAN-2001; 2001US-262164P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-263131P.
 PR 17-JAN-2001; 2001US-263599P.
 PR 19-JAN-2001; 2001US-262760P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263066P.
 PR 19-JAN-2001; 2001US-263069P.
 PR 19-JAN-2001; 2001US-263070P.
 PR 19-JAN-2001; 2001US-263074P.
 PR 19-JAN-2001; 2001US-263076P.
 PR 19-JAN-2001; 2001US-263077P.
 PR 19-JAN-2001; 2001US-263329P.
 XX
 XX (INCYTE GENOMICS INC.
 XX Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
 PI Panzer SR, Lin TE, Harris B, Flores V, Dafio A, Marwaha R, Chen AJ;
 PI Chang SC, Gerstin EH, Petalita CH, David MH, Lewis SA;
 XX WPI; 2002-590716/63.
 DR P-PDB; ABG69816.
 XX
 XX New purified secretory polypeptides and polynucleotides, useful in the
 PT diagnosis, study, prevention or treatment of diseases associated with
 PT decreased expression of functional secretory molecules, e.g. AIDS,
 cancer or allergies

PT amplicons from chemically treated DNA -
XX

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW

XX Oligonucleotide for detecting cytosine methylation by the

KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX Homo sapiens.
 OS
 PN WO200218632-A2.
 XX
 XX 07-MAR-2002.
 PD
 PF 01-SEP-2001; 2001WO-EP10074.
 XX
 XX 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 PA
 PI Olek A. Piepenbrock C, Berlin K, Guetig D;
 DR WPI; 2002-371829/40.
 XX
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA
 XX
 PS Claim 12; 56pp + Sequence listing; 56pp; German.
 XX
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridization to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridized to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC systems etc., particularly for detecting mutations or single nucleotide
 CC polymorphisms (SNPs); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 XX Sequence 799 BP; 289 A; 297 C; 95 G; 118 T; 0 other;
 SQ
 Query Match 11.1%; Score 343.8; DB 24; Length 799;
 Best Local Similarity 83.5%; Pred. No. 1.7e-75;
 Matches 390; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 387 AGTAAGGAGAGAGGCGCGGCTTCGCGCTCCCGGCGATACGATGCTGACGCTCCG 446
 Db 498 AGTAAGGAGAGAGGCGCGGCTTCGCGCTCCCGGCGATACGATGCTGACGCTCCG 439
 QY 447 TCGGCGTGGCTGAGAGGAGAGGCGCGGCGCGCGCGATGCTGATTTCCG 506
 Db 438 TCGGCTGGCTGAGAGGAGAGGCGCGGCGCGCGCGATGCTGATTTCCG 379
 QY 507 GTCCGACAGTGGCTGAGAGGAGAGGCGCGGCGCGCGCGATGCTGATTTCCG 553
 Db 378 GTCCGACAGTGGCTGAGAGGAGAGGCGCGGCGCGCGCGATGCTGATTTCCG 332
 RESULT 9
 ABV15331/C
 ID ABV15331 standard; cDNA; 356 BP.
 XX
 AC ABV15331;
 XX
 DT 13-SEP-2002 (first entry)
 DE
 XX Human prostate expression marker cDNA 15322.
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200160860-A2.
 PN
 XX 23-AUG-2001.
 PD
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 XX 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI
 PI Schlegel R, Endege WO, Monahan JE;
 DR WPI; 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX
 PS Claim 1; Page 2572; 11750pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 356 BP; 97 A; 80 C; 85 G; 94 T; 0 other;
 Query Match 9.4%; Score 292.4; DB 23; Length 356;

CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV6213) of the
CC	specification or its complement. (I) is useful for:
CC	(a) assessing whether a patient is afflicted with prostate cancer;
CC	(b) monitoring the progression of prostate cancer in a patient;
CC	(c) assessing the efficacy of a test compound to inhibit prostate
CC	cancer in a patient;
CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC	in a patient;
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;
CC	(f) assessing the prostate cell carcinogenic potential of a compound;
CC	(g) determining whether prostate cancer has metastasized in a patient;
CC	(h) assessing the aggressiveness or indolence of prostate cancer in a
CC	patient;
XX	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX	
SQ	Sequence 439 BP; 113 A; 104 C; 110 G; 112 T; 0 other;
Query Match:	9.3%; Score 288.8; DB 23; Length 439;
Best Local Similarity	91.4%; Pred. No. 6,8e-62;
Matches 338; Conservative	0; Mismatches 7; Indels 25; Gaps 2;
Db	
Oy	1687 GATTAACAAAAGCTCTTTGGAACTCATCTCATTGCAAGACATGTGAAGAATGCCTTCGG 1746
Db	439 GATAACA AAAAGCTTGTGGAACTCATCTCATTGCAAGACATGTGAAGAATGCCTTCGG 380
Oy	1747 TCCTTTTGACAGTCTCTCGAGTGGCTTTCCCCAAGTCTACTATATATATATACCAGG 1866
Db	379 TCCTTTTGACAG--TCTCTCGAGTGGCTTTCCCCAAGTCTACTATATATATACCAGG 322
Oy	1807 CTAAAGCATGTATGACATCTCTTTGTTACTCATCTGTTCCGCCATCTGTAGTCTTATTC 1866
Db	321 CTAAAGCATGTATGACATCTCTTTGTTACTCATCTGTTCCGCCATCTGTAGTCTTATTC 262
Oy	1867 AGATTCATGACATPAACAAGGCGTGCACAGAGAGATPAAGCTTGTCATATCTTAGAGAT 1926
Db	261 AGATTCATGACATPAACAAGGCGTGCACAGAGAGATPAAGCTTGTCATATCTTAGAGAT 202
Oy	1927 TTGCCACCTTGCAGATGAGTTTATGACATTTTATTAATPAGGAGAGAGGTTGATGC 1986
Db	201 TTGCCACCTTGCAGATG-------AGGCAGAGAAAGTTGATGC 165
Oy	1987 AGGCGTTCAACCATGCTGTTGAACAGAACCCCAAAGGCAACATTTGGCTGTTTAGG 2046
Db	164 AGGCGTTCAACCATGCTGTTGAACAGAACCCCAAAGGCAACATTTGGCTGTTTAGG 105
Oy	2047 TACCTGGGTC 2056
Db	104 TACCTGGGC 95
RESULT 11	
ABV45182/C	
ID	ABV45182 standard; CDNA: 439 BP.
XX	
AC	ABV45182;
DT	16-SEP-2002 (first entry)
DE	
XX	Human prostate expression marker CDNA 45173.
XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM	pharmacogenomic marker; gene; ss.
OS	Homo sapiens.
XX	
PN	WO200160860-A2.
PD	
PF	20-FEB-2001; 2001WO-US05171.
FR	17-FEB-2000; 2000US-183319P-
	16-MAR-2000; 2000US-189662P.

PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 PA (MILL.) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 DR Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX
 PS Claim 1; Page 8947; 11750pp; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 439 BP; 113 A; 104 C; 110 G; 112 T; 0 other;
 XX
 Query Match 9.3%; Score 288.8; DB 23; Length 439;
 Best Local Similarity 91.4%; Pred. No. 6.8e-62; Indels 25; Gaps 2;
 Matches 338; Conservative 0; Mismatches 7;
 QY 1687 GATACAAAAGGCTTTGGAGCTCATCTCATGCAAGACATGGTGAAGATGCACCTTGG 1746
 DB 439 GATACAAAAGGCTTTGGAGCTCATCTCATGCAAGACATGGTGAAGATGCACCTTGG 380
 QY 1747 TCTTTGTGAGATCTCCGAGTGTTCCTCCCAAGCTCACTATTAATACACACAG 1806
 DB 379 TCTTTGTGAG--TCTCCGAGTGTTCCTCCCAAGCTCACTATTAATACACACAG 322
 QY 1807 CTAGGAGCTGATGAGCTCTTGTCTACTAGTGTTCGGCAATTCGTAGTCTTATTTC 1866
 DB 321 CTAGGAGCTGATGAGCTCTTGTCTACTAGTGTTCGGCAATTCGTAGTCTTATTTC 262
 QY 1867 AGATCCATGACATPAACAGGCTCTGAACAGAGAGTAAGCTTGCAATCTTGAGGAAT 1926
 DB 261 AGATCCATGACATPAACAGGCTCTGAACAGAGAGTAAGCTTGCAATCTTGAGGAAT 202
 QY 1927 TTGGCACTTGCAGAGTGAATTAATGACATTTATTAATAGGACAGAGGTTGATGC 1986
 DB 201 TTGGCACTTGCAGAGTGAATTAATGACATTTATTAATAGGACAGAGGTTGATGC 165
 QY 1987 AGCGCTTCAACCATGCTGTGTAACAGAGAACCCCAAGGCAACATTTGGCTCTTTAGG 2046
 DB 164 AGCGCTTCAACCATGCTGTGTAACAGAGAACCCCAAGGCAACATTTGGCTCTTTAGG 105
 QY 2047 TACCTGGGTC 2056
 DB 104 TACCTGGGTC 95
 AC
 RESULT 12
 ID AB051660 standard; DNA; 799 BP.
 XX AB051660;

XX 12-JUL-2002 (first entry)
 DT Oligonucleotide for detecting cytosine methylation SEQ ID NO 38251.
 XX
 DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 XX drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200216632-A2.
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-BP10074.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 PA (EPIG-) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX WPI; 2002-371829/40.
 DR
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one
 CC member of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridization to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridized to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cellular differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 SQ Sequence 799 BP; 122 A; 95 C; 262 G; 320 T; 0 other;
 XX
 Query Match 8.9%; Score 274.2; DB 24; Length 799;
 Best Local Similarity 73.9%; Pred. No. 4.1e-58;
 Matches 348; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
 QY 86 CTGCTGCTTGTGCTGCAACCACTAAGCTTACGCAAACTTCACGCTTCTTCCGCC 145
 DB 799 CTACTGCTTGTGCTGCAACCACTAAGCTTACGCAAACTTCACGCTTCTTCCGCC 740
 QY 146 TTGGCGTCACTTTCTTAAGAAATTTCCAGAGGCGAGGCAAGAGGCGGCTCTGAGAC 205
 DB 739 TTGGCGTCACTTTCTTAAGAAATTTCCAGAGGCGAGGCAAGAGGCGGCTCTGAGAC 680
 QY 206 TCCGGGCTCCGCTCTTTCGAGAACCGCCCACTACCCAGACTCCGACAGAGGGTGAAC 255
 DB 679 TCCGAATCCGCGCTCTTTCGAGAACCGCCCACTACCCAGACTCCGACAGAGGGTGAAC 620
 QY 266 AAGATTAATCTTCGGGTTCGGATGCTCTTAATCTCGGAGAAAGAGAGGCGCGCA 325

Db	619	AAAAAATTA	CTTCGAT	CTCGAT	CTCTCT	CTTA	TA	CTTCG	CGAAAA	AAAAA	AAAAA	ACG	CGCC	560
Qy	326	TCGCGCGAA	CGAGCG	CGTGCG	CGAGG	GGG	G	TGG	CGCGG	AGG	CGG	AG	CGCC	385
Db	559	TCATCCGAA	CGAAAA	GATTA	ACGAAAA	AAAAA	TTA	ACGAAAA	ACG	CGAA	ATCC	CGA	500	
Qy	386	GAGTAAG	GAGAG	GGGG	CGGG	G	TGCG	CGTCC	CGG	CA	TAC	GCAT	CGCT	445
Db	499	AAATTA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	440
Qy	446	GTCGGG	CTGG	CTGAG	AGG	GGG	CGG	CGG	CGG	CGG	CGG	CGG	CGG	505
Db	439	ATCGAA	CTAA	ACTTA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	380
Qy	506	CGTCGG	AGCG	GTGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	CGCG	556
Db	379	AATCGAA	CA	TA	AGT	TA	ACG	CGCG	CGA	TA	TA	ACCG	CGAAAA	329

RESULT 13
ABQ51661
ID ABQ51661 standard; DNA; 799 BP.

AC ABQ51661;
XX
DT 12-JUL-2002 (first entry)

Oligonucleotide for detecting cytosine methylation SEQ ID NO 38252

XX Human; cytosine methylation; 5'-CPG-3'; uracil; cytosine; diagnosis;
KM drug; side effect; cancer; central nervous system; cardiovascular;
KM gastrointestinal; respiratory system; single nucleotide polymorphism
KM SNP; cell differentiation; ds.

OS	Homo sapiens.
XX	
PN	WO200218632-A2.

PD	07-MAR-2002.
XX	
PF	01-SEP-2001; 2001WO-EP10074.

PR	01-SEP-2000; 2000DE-10
PR	05-SEP-2000; 2000DE-10
XX	
PA	(EPIG-) EPIGENOMICS AG

PI Olek A, Piepenbrock C, Berlin K, Gnetig D
XX
DR WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
PS Claim 12; 56pp + Sequence listing; 56pp; German.

CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ1310-ABQ4121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.

Sequence 799 BP; 320 A; 262 C; 95 G; 122 T; 0 other

Query Match	8.9%	Score 274.2;	DB 24;	Length 799;
Best Local Similarity	73.9%	Pred. No. 4.1e-58;		
Matches 348;	Conservative	0;	Mismatches 123;	Indels 0;
				Gaps 0

Qy	86	CTGCTCGTTGCTGCTGCTGCAACCACTAAGTCTAGCGAAACCTCCACGGTTTCCTTCCGCC	145
Dp	1	CTACTCGTTGCTGTACTCGCAACCACTAAATCTACGCAAACTTCCAGATTTCTTCGGCC	60
Qy	146	TTCCGGGTCAACCTTTTAAAGAAATTCGCCAGAGGCGCAGACGGGGCGGCGTCTGAGAC	205
Dp	61	TTCCGGGTCAACCTTTTAAABAAATTCGCAAAAACCAACCAACGAAACCTCTPAAAC	120
Qy	206	TCCGGGCTCCGCTCTTTTCCGAGAAACCGGCCACTACCCAGACTCCGACAGAGGGTGAAA	265
Dp	121	TCCGAACCTCCGCTCTTTCCGAAAAACCGGCCACTACCCAAAACTCCGACAAAAATTA	180
Qy	266	AAAGATTAATTCGCGCTCGGATGCTCTAATCTCGGAGAGAGAGCGCGCGCA	325
Dp	181	AAAAATTAATTCGATCTCGCATCTCTCTAATCTCGGAAAAAAAACGACGCGCA	240
Qy	326	TCGCGCGAAACGAGCGGTGCGGAGGAGGGGTGTGCGCGGAGACCGGAATCCCGG	385
Dp	241	TGCACCGAAACGAAACGATTAACGAAAAAAAATATTAACGAAAAACGGGAATCCCGA	300
Qy	386	GAGTAAGGAGAGAGGGGGCGGCTCGCGCTCCGGGCGATACGATCGTGCACGCTGCC	445
Dp	301	AAATTAACGAAAAAAACGAAATCGGCGCTCCGAAACATACGATACGATACGATACG	360
Qy	446	GTCGGGCTGGCTGAGAGGAGGAGGGGCGCGCGGCCGAGCGCGTCTTATTTTCGT	505
Dp	361	ATCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	420
Qy	506	GATCGGACAGTGTGTGCGCGCGCGGCTGACACGGGAGAAAGTACGATTA	556
Dp	421	AATCGAACAATAGTAAACGACGGAATTAACGAGAAAAATTAATTAATTAATTAATTAAT	471

RESULT 14
ABV06162/c
ID ABV06162 standard; cDNA; 303 BP.

AC	ABV06162;
XX	
DT	13-SEP-2002 (first entry)

XX
DE Human prostate expression marker cDNA 6153.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker
KM pharmacogenomic marker; gene; ss.
KW

xx
OS Homo sapiens

PN WO200160860-A2.
XX
PD 23-AUG-2001.

AA 20-FEB-2001; 2001WO-US05171.
PF

XX	17-FEB-2000;	2000US-183319P
PR	16-MAR-2000;	2000US-189862P
PR	25-MAY-2000;	2000US-207454P
PR	09-JUN-2000;	2000US-211314P
PR	18-JUL-2000;	2000US-219007P
PR	13-DEC-2000;	2000US-255281P

QY 1119 GAGACATC---TTGAGTTACAGCATGCTAAAGATGTGAGAGATGACATGCAAGAG 1175
 DB 830 GGCACAGCCAGCCAGCTTTATAGCTCGCTAAAGAGATGTGAGAGATGACATG--- 886
 QY 1176 AGTAAAGATGACTGAGATGACAGAGAGAGAGAGATCCAGAAATTGAACTAGAGAA 1235
 DB 887 -----CAAGTGCMAAAAGCTGACTTCACGAGAGAGCTGATG 924
 QY 1236 CCAACATGTTTACAGATTTTACAGAGATGAAATTTACTGTTTACTGACAGTGT 1295
 DB 925 CCGGTGTTACCGGCTGCTT---TCATGGCCATCTTTTCAGATGATCATCAGCTGAG 982
 QY 1296 TATAGCTTTACTAAGAAAGAGACCAAGTGTGTCAGAACTCMAAAATTTAGTTCA 1355
 DB 983 GCAATGGCCAGCATATGATCCGTCGATGCAAGGTAGACATATACMAATCTGTTGT 1042
 QY 1356 AGCAGCATCTTCTTCTGCAATTCATTAATTCATTCATTCATTCATTCATTCATTC 1415
 DB 1043 GGCCTCCGAGATGCTTCCAGAGATTAAGAGAACTTTGATCGGGGACACACACAGAA 1102
 QY 1416 TGATACTACAAAGAGATCATCCATTAATGATGAGTTTGAACCCCTTGTGAACAG 1475
 DB 1103 GGGGTCT-----GACGCTCCCAATCCATGGTTTCTCTCCGCGCATCCAGACCG 1153
 QY 1476 GCTACTTCCACTTACTCTCCCTGATATGCAAAATTAATAAGGAGAAATGTGAA 1535
 DB 1154 CAGCCAAACCCGCGCTTCCGCTGATGATTAAGATTAAGATTAAGATTAAGATTA 1213
 QY 1536 CTATTTGCAAGATTAATAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1595
 DB 1214 GTTCTAGAGAAATGATTTGCGCTTCAAAATAGCCTGCAAAAGTCAACAGTAAGA 1273
 QY 1586 TTTACATTTGATCTGATTTTCTGTAATTAAGAACAGTCT---ACCATGTTCT 1552
 DB 1274 TTACTATTCGGCGCTGAACTTCTTTATGAGTACGCAAAAGTCCGGGCTGATCTCT 1333
 QY 1653 TTCAAGATCTCTGTTACAAACCACTTCTGATGAGATTAACAAAGATCTTTGAACTCA 1712
 DB 1334 GTCCAGAGAGCTGTGCAAAACCTGTTGAGCGCAACATGCGTATGAGGACGAGAAAGT 1393
 QY 1713 TCTCATGCAAGATGATGAGAAATGATGATGATGATGATGATGATGATGATGAT 1772
 DB 1394 TCCCATTAAGAGATTTCTGCGCCACTCGGTTCAAGTCTGCAACCTGCAATTTGAA 1453
 QY 1773 TCCCCAGAGTGTACTATTAATTAATGATGATGATGATGATGATGATGATGAT 1832
 DB 1454 T---GCCAAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
 QY 1833 ACTCATGTTT-----CGGCAATTTCTGATCTTATTCAGATGATGATGATGATGAT 1886
 DB 1512 CTTACTAGTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 1571
 QY 1887 GGTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1946
 DB 1572 GGCACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1631
 QY 1947 TTTATGACATTTTATTTATTAATGAGCAAGAGATGATGATGATGATGATGATGAT 2006
 DB 1632 GAGAGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1691
 QY 2007 TGAACAGAGAACCCCAAG--GCAACATTTGAGCTGTTTATGATGATGATGATGAT 2065
 DB 1692 GAGCCATGAGCAGCTGATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1751
 QY 2066 AACCTTGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2125
 DB 1752 TGTTCCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1811
 QY 2126 CACGAGTACTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2185
 DB 1812 CATGATTTCTTCAATCTAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1871
 QY 2186 ACATTGATGCTGCGATGCTCTCAATGCGACAGAGAAAGATTAATGAGAGACAGCAG 2245

DB 1872 GCGCTGACGCGCAAGAGACATCTCTCTGCGCCAGAGAGATGACCGGACACCAAGT 1931
 QY 2246 AAAGCCCTAGTAG-----TAAAAAAACAAG 2272
 DB 1932 AAGACCTAGTGGGGGCGAGCGAGCGGCTAAAAACCGAAGCGCGCCAGCCGAAAAAG 1991
 QY 2273 AAAAAAAGAAAGTTCCGCTTGAAGCCAGAGATGACATGAGCAGATATGAGAAC 2332
 DB 1992 AACAAAAAGACTCAAAAGGCGTACCGCGCAGATGCTTCTTACACGCGCTACTGAC 2051
 QY 2333 ATGTGCTGAGATGTTTAAACCATGATGATGATGATGATGATGATGATGATGATGATGAT 2392
 DB 2052 CTCTGCGCGAGATGATTAAGAGATGAGCGCGCTGACCAAAAGAGAGAGGCTGACTG 2111
 QY 2393 CCGAAGTTGAGCTTGAATGAGAACAGATGCGTATGACACAGAGTTGCTCATTCAG 2452
 DB 2112 CCGCTGTCMAAGTTGAGACAAAGAGAGATCCGCTACAAACCGAGCTTCTGCTTCCG 2171
 QY 2453 AGTGTGATGACCCGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2512
 DB 2172 AACCTACAGATCTCCGCGATTTGATGAGAGATTTCAAAACGTCGAGAGACATG 2231
 QY 2513 AATATAGCCCTCTCTCTGATGCTCTGATGATGATGATGATGATGATGATGATGAT 2572
 DB 2232 A-----TGCAATAGCGCTGAGATCTATTAATGATGATGATGATGATGATGAT 2282
 QY 2573 CAGGCAAAATGATTAATGAAATATTTCTTAACCGGACATGATGATTAATGAAATTTTA 2632
 DB 2283 CAGGCGCGAATGTTTCAAAAGCATTTCAAAACCATGATGATGATGATGATGATGAT 2342
 QY 2633 AAGGTGCCAAACCAACTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2692
 DB 2343 CAATGACAGACATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2402
 QY 2693 TCTAAGTTCTCTGAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2751
 DB 2403 GTCAAGCGGAGCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2461

Search completed: November 24, 2003, 22:01:07
 Job time : 790 secs

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Pred. No. is the number

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	length	DB	ID	Description
1	3095	100.0	3095	6	AX533454	AX533454 Sequence
2	2477.2	78.1	2488	6	AX537452	AX537452 Sequence
3	1838.4	59.4	2552	10	AY102701	AY102701 Mus muscu
4	1799.2	58.1	2543	10	AF72892	AF72892 Rattus no
5	1766.6	57.1	1839	9	AK055286	AK055286 Homo sapi
6	1264.8	40.9	1457	9	BC053286	BC053286 Danto rer
7	988	31.9	997	9	AK026226	AK026226 Homo sapi
8	885.4	28.6	1796	9	AK056059	AK056059 Homo sapi
9	717.4	23.2	719	6	AX664135	AX664135 Sequence
10	664.2	21.5	147876	6	AC102010	AC102010 Mus muscu
11	551.4	17.8	181864	9	AL133743	AL133743 Human DNA
12	398.4	12.9	160115	9	AL151447	AL151447 Human DNA
13	321.4	10.4	446	9	HSAB334621	AJ334482 Homo sapi
14	299	9.7	743	9	HSAB334413	AJ334431 Homo sapi
15	236.4	7.6	642	9	HSAB442338	Homo sapi
16	236.4	7.6	742	9	HSAB31811	AJ331811 Homo sapi
17	233.2	7.5	640	9	HSAB323329	AJ342329 Homo sapi
18	233.2	7.5	652	9	HSAB12337	AJ342327 Homo sapi
19	217	7.0	2537	3	AY060948	AY060948 Drosophil
20	207.6	6.7	2565	3	AY089452	AY089452 Drosophil
21	207	6.7	320902	11	AL161456	AL161456 Homo sapi
22	182	5.9	489	11	G63197	G63197 SHGC-140830
23	177.8	5.7	34935	9	AL161453	AL161453 Human DNA
24	177.6	5.7	111554	2	AC143812	AC143812 Mecaaca mu
25	155.8	5.0	113836	10	AL1669837	AL1669837 Mouse DNA
26	153.2	4.9	253611	2	AC111473	AC111473 Rattus no
27	153.2	4.9	271861	2	AC108541	AC108541 Rattus no
28	153.2	4.9	281804	2	AC1334869	AC1334869 Mus muscu
29	140.4	4.5	151	6	AX523453	AX523453 Sequence
30	130.4	4.3	204131	6	AC128898	AC128898 Rattus no
31	130.4	4.2	3497	10	AY102702	AY102702 Mus muscu
32	130.4	4.2	269131	2	AC135469	AC135469 Mus muscu
33	90.6	2.9	16016	2	AC020508	AC020508 Drosophil
34	90.6	2.9	60135	2	AC091200	AC091200 Drosophil
35	90.6	2.9	164443	3	AC012167	AC012167 Drosophil
36	90.6	2.9	298827	3	AE003463	AE003463 Drosophil
37	89.4	2.8	815	6	AX23523	AX23523 Sequence
38	67.8	2.2	7218	6	164494	164494 Sequence 14
39	66.6	2.2	2000	6	AX655593	AX655593 Sequence
40	66.6	2.2	110737	2	AC011105	AC011105 Homo sapi
41	65.2	2.1	159980	2	AP005743	AP005743 Oryza sat
42	65.6	2.1	68328	2	AC137705	AC137705 Homo sapi
43	64.6	2.1	932	11	BY026280	BY026280 S212P647
44	64.4	2.1	69375	2	AC101456	AC101456 Mus muscu
45	64	2.1	67769	2	AC068336	AC068336 Homo sapi

ALIGNMENTS

RESULT 1				
AX523454				
LOCUS	AX523454	3096 bp	DNA	linear
DEFINITION	Sequence 42 from Patent WO02064788.			
ACCESSION	AX523454			
VERSION	AX523454.1	GI:24412350		
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1			
AUTHORS	Mecina, R. A., Recipon, H., Chen, S. Y., Sun, Y. and Liu, C.			
TITLE	Compositions and methods relating to lung specific genes and proteins			

QY 1981 TGATGAGGCGCTTACACCACTGCTGTTGAAACAGAACCCCAAGGCAACATTTGGCCCTG 2040
 DB 1981 TGATGAGGCGCTTACACCACTGCTGTTGAAACAGAACCCCAAGGCAACATTTGGCCCTG 2040
 QY 2041 TTTAGTACTGCGTCTTACCATACCTTGCATTTATGATACAGTACCTTCTAAGTGG 2100
 DB 2041 TTTAGTACTGCGTCTTACCATACCTTGCATTTATGATACAGTACCTTCTAAGTGG 2100
 QY 2101 CTTTGAATGGAATCTACAGATGCAAGATCTATTTACATATTTGGTATCTCTCTA 2160
 DB 2101 CTTTGAATGGAATCTACAGATGCAAGATCTATTTACATATTTGGTATCTCTCTA 2160
 QY 2161 ATTCTTTAGCGATGTTGATGTCATCATTTAGTCGCGCGCATGCTCTCAATGCGAGA 2220
 DB 2161 ATTCTTTAGCGATGTTGATGTCATCATTTAGTCGCGCGCATGCTCTCAATGCGAGA 2220
 QY 2221 GGAAAGGATTAATGGAAGAGCAGCAAGAAAGCGCGTAGTAGTAAAAAAGAAAAA 2280
 DB 2221 GGAAAGGATTAATGGAAGAGCAGCAAGAAAGCGCGTAGTAGTAAAAAAGAAAAA 2280
 QY 2281 GAAAGTTCCCGCATTTGAGCCGAGATGCAATGAGCCAGCATATGCAACATGTCG 2340
 DB 2281 GAAAGTTCCCGCATTTGAGCCGAGATGCAATGAGCCAGCATATGCAACATGTCG 2340
 QY 2341 TGGAAATGTTAAACCATGTAGCATTTGACATGAGCGCAAGTACGTAAACCGAAGTT 2400
 DB 2341 TGGAAATGTTAAACCATGTAGCATTTGACATGAGCGCAAGTACGTAAACCGAAGTT 2400
 QY 2401 TGAGCTTGTAGTGAACAAATGTTGGTATGAAACAGAGTTGCTCCATTCACAGTGTAT 2460
 DB 2401 TGAGCTTGTAGTGAACAAATGTTGGTATGAAACAGAGTTGCTCCATTCACAGTGTAT 2460
 QY 2461 GACCCCGCGCGAGTCACTACTACAGTCAAGGAAATGTCGATCAATTAATATAG 2520
 DB 2461 GACCCCGCGCGAGTCACTACTACAGTCAAGGAAATGTCGATCAATTAATATAG 2520
 QY 2521 CCTCTCTCTCACTCTCTGAACTGATGATGTCAGTATGATGCACTTCAACAGGCAA 2580
 DB 2521 CCTCTCTCTCACTCTCTGAACTGATGATGTCAGTATGATGCACTTCAACAGGCAA 2580
 QY 2581 AATGATATGGAATAATTCCTTAACCGGACATGAGTTAATAGATTTTAAAGGTTGC 2640
 DB 2581 AATGATATGGAATAATTCCTTAACCGGACATGAGTTAATAGATTTTAAAGGTTGC 2640
 QY 2641 CAATCCCACTTTGTGTTATGAAATGATGAGGAGGACAAAAAGGAATCTAAAGT 2700
 DB 2641 CAATCCCACTTTGTGTTATGAAATGATGAGGAGGACAAAAAGGAATCTAAAGT 2700
 QY 2701 TCTCTCTGAAATTTGATTTCTCTGCTCATATAATTTCTGTTGTAACCTGTTTGA 2760
 DB 2701 TCTCTCTGAAATTTGATTTCTCTGCTCATATAATTTCTGTTGTAACCTGTTTGA 2760
 QY 2761 GAGACTGGGAGGTGCGCATTAAGGGGAGAGTCTTTTCAACCCAACTCTTAAGG 2820
 DB 2761 GAGACTGGGAGGTGCGCATTAAGGGGAGAGTCTTTTCAACCCAACTCTTAAGG 2820
 QY 2821 CACATCACCAAGCTCCACATGACGGAAGTGAATGATTTCTGGGTAACACTCATTA 2880
 DB 2821 CACATCACCAAGCTCCACATGACGGAAGTGAATGATTTCTGGGTAACACTCATTA 2880
 QY 2881 TAAGGAATACCTTTAGTTGACAGCTTATATGACATGATGAAACTGCTGTTTAA 2940
 DB 2881 TAAGGAATACCTTTAGTTGACAGCTTATATGACATGATGAAACTGCTGTTTAA 2940
 QY 2941 TGGTTATATATGTTTCATGGAAGAACTGCTTATTAATGATGATGATGATGAT 3000
 DB 2941 TGGTTATATATGTTTCATGGAAGAACTGCTTATTAATGATGATGATGATGAT 3000
 QY 3001 GGTTTTATACAGATTTAATCAAAATCAATTTTATGATGATGATGATGATGAT 3060
 DB 3001 GGTTTTATACAGATTTAATCAAAATCAATTTTATGATGATGATGATGATGAT 3060
 QY 3061 TTATTAAGGTTAATTAATTTCTTGACAAAAA 3096

DB 3061 TTATTAAGGTTAATTAATTTCTTGACAAAAA 3096
 RESULT 2
 Locus AX537452
 Definition Sequence 23 from Patent WO020709.
 Accession AX537452
 Version AX537452.1 GI:25269238
 Keywords
 Source
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS
 1 Lal, P.G., Baughn, M.R., Yao, M.G., Walla, N.K., Elliot, V.S., Xu, Y.,
 Honchell, C.D., Yue, H., Ding, L., Gietzen, K.J., Ison, C.H., Lu, D.A.,
 Hafalia, A.J., Ghandi, A.R., Thangavelu, K., Sanjwal, M.,
 Tang, Y.T., Ramkumar, J., Griffin, J.A., Swarnaker, A., Azimzai, Y.,
 Saperstein, S.K., Burford, N., Lee, E.A., Lu, Y., Tran, U.K. and
 Marquis, J.P.
 Molecules for disease detection and treatment
 Patent: WO 020709-A 23 12-SEP-2002;
 Incyte Genomics, Inc. (US)
 Location/Qualifiers
 1..2488
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 711920CB1"
 BASE COUNT 768 a 461 c 567 g 692 t
 ORIGIN
 Query Match 78.1%; Score 2417.2; DB 6; Length 2488;
 Best local Similarity 98.5%; Pred. No. 0;
 Matches 2470; Conservative 0; Mismatches 13; Indels 25; Gaps 2;
 QY 474 GCGCGCGCGCGAGCGCGGCGTGTATTTCCGTGTCGCGACAGTCGCGCGCGCGGT 533
 DB 1 GCGCGCGCGCGAGCGCGGCGTGTATTTCCGTGTCGCGACAGTCGCGCGCGCGGT 60
 QY 534 GACCAAGGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 593
 DB 61 GACCAAGGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 594 ATGGAGCTCAGTATGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 653
 DB 121 ATGGAGCTCAGTATGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 654 CCAAGATTTTGAAGAGCTTTCGAGATTTAAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 713
 DB 181 CCAAGATTTTGAAGAGCTTTCGAGATTTAAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 240
 QY 714 ATTGGCTCTTTTGAAG 773
 DB 241 ATTGGCTCTTTTGAAG 300
 QY 774 CATGATGGAACCAAGATTAATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 833
 DB 301 CATGATGGAACCAAGATTAATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 834 CAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 893
 DB 361 CAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
 QY 894 TTGCTGTTGATTAACGTTGTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 953
 DB 421 TTGCTGTTGATTAACGTTGTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 954 TTACATTCATTAATCCAGACTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1013
 DB 481 TTACATTCATTAATCCAGACTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 1014 CTTGAAAATCTGTGACATTCGACAGGAAAAAGTAAATTAAGCTGCTGTTTGGAGAGA 1073
 Db 541 CTTGAAAATCTGTGACATTCGACAGGAAAAAGTAAATTAAGCTGCTGTTTGGAGAGA 600
 QY 1074 AGATTTTCAGTCAATGACTTATGATTTAAATGCGCTAATGAGTGAAGATCTTCAGT 1133
 Db 601 AGATTTTCAGTCAATGACTTATGATTTAAATGCGCTAATGAGTGAAGATCTTCAGT 660
 QY 1134 TACAGGCACTGTAATGAAGTGTGAGATGACATGCAAGAAAGTAAGATGACTGAG 1193
 Db 661 TACAGGCACTGTAATGAAGTGTGAGATGACATGCAAGAAAGTAAGATGACTGAG 720
 QY 1194 TCGACAGAGAGAGAGAGAGATCCAGAACTTGAATCTGAACACCAACAAAGTTAGCAGT 1253
 Db 721 TCGACAGAGAGAGAGAGATCCAGAACTTGAATCTGAACACCAACAAAGTTAGCAGT 780
 QY 1254 ATTGCGAGAGTGAATTTAATGCTGTGTTACTGACAGTGTATAGCTTTACTTAAGAA 1313
 Db 781 ATTGCGAGAGTGAATTTAATGCTGTGTTACTGACAGTGTATAGCTTTACTTAAGAA 840
 QY 1314 AGAGACAGAGTGTGACAGAGCTGAAAATGATGTTGTTCAAGAGAGATCTTCTTC 1373
 Db 841 AGAGACAGAGTGTGACAGAGCTGAAAATGATGTTGTTCAAGAGAGATCTTCTTC 900
 QY 1374 TGGCATTCAATTCCTTTCATCATGAGGATCCAGGCGCAGATGATACAAAAGAGA 1433
 Db 901 TGGCATTCAATTCCTTTCATCATGAGGATCCAGGCGCAGATGATACAAAAGAGA 960
 QY 1434 TCATCCAAATGATGAGGTTTGAACCCCTGTGAACAGAGCTACTCCACTACCTT 1493
 Db 961 TCATCCAAATGATGAGGTTTGAACCCCTGTGAACAGAGCTACTCCACTACCTT 1020
 QY 1494 CCGCTGATGCAAAAAATTAATTAAGGAGAGAAATGATGATATTTTTCAGATTAAT 1553
 Db 1021 CCGCTGATGCAAAAAATTAATTAAGGAGAGAAATGATGATATTTTTCAGATTAAT 1080
 QY 1554 AGATGAGATTAATACTGTCTGTGAGGTTGATTTAACAATTTACATTTGATGCTGGA 1613
 Db 1081 AGATGAGATTAATACTGTCTGTGAGGTTGATTTAACAATTTACATTTGATGCTGGA 1140
 QY 1614 TTTTTCGTGAATTTAGTGAACAGTCACTGCTTTCAAGATCTGTTTACAAC 1673
 Db 1141 TTTTTCGTGAATTTAGTGAACAGTCACTGCTTTCAAGATCTGTTTACAAC 1200
 QY 1674 CACTTCTGTGATTAACAAAAGGCTTTGGAATCTATCTATGCAAGCATGCTGA 1733
 Db 1201 CACTTCTGTGATTAACAAAAGGCTTTGGAATCTATCTATGCAAGCATGCTGA 1280
 QY 1734 AGATGCACTTGGCTTTTGTGACATCTCCGAGTGTTCGCCCAAGTGTACTATAT 1793
 Db 1261 AGATGCACTTGGCTTTTGTGACATCTCCGAGTGTTCGCCCAAGTGTACTATAT 1318
 QY 1794 AATTAACACAGGCTPAGACTGTATGCACTCTTTTATCTCACTGTGTTGGCATTG 1853
 Db 1319 AATTAACACAGGCTPAGACTGTATGCACTCTTTTATCTCACTGTGTTGGCATTG 1378
 QY 1854 TGTAGCTTATTCAGATTCATGACATTAACAGGAGCTGACAGAGAGATTAAGCTGAT 1913
 Db 1379 TGTAGCTTATTCAGATTCATGACATTAACAGGAGCTGACAGAGAGATTAAGCTGAT 1438
 QY 1914 ATTCTTGAGGAATTTGCACTTGCAGAGATGAGTTATGACATTTATTTAATAGGAG 1973
 Db 1439 ATTCTTGAGGAATTTGCACTTGCAGAGATGAGTTATGACATTTATTTAATAGGAG 1475
 QY 1974 AGAAGGTTGACAGGCTTCAACCATGCTGTGTAAGAGAGAACCCCAAGCAAT 2033
 Db 1476 AGAAGGTTGACAGGCTTCAACCATGCTGTGTAAGAGAGAACCCCAAGCAAT 1535
 QY 2034 TGGGCTGTTTGAAGTCTGGGCTCTTACCATTAACCTTGCATTTATGATCACTTTC 2093
 Db 1536 TGGGCTGTTTGAAGTCTGGGCTCTTACCATTAACCTTGCATTTATGATCACTTTC 1595

QY 2094 TAACTGCTTTGAATTTGGAATCTACAGTATGACAGAGTACTATTAATATTTGATC 2153
 Db 1596 TAACTGCTTTGAATTTGGAATCTACAGTATGACAGAGTACTATTAATATTTGATC 1655
 QY 2154 TCTGTGAATCTTTACAGTATGATGATGATCAATTTGAGTGTGCTGCTCAAA 2213
 Db 1656 TCTGTGAATCTTTACAGTATGATGATGATCAATTTGAGTGTGCTGCTCAAA 1715
 QY 2214 TGGCAGAGAGAGATTAATGGAAGAGCAGAGAAAGGCGTATGATTAATAAAACAAAGA 2273
 Db 1716 TGGCAGAGAGAGATTAATGGAAGAGCAGAGAAAGGCGTATGATTAATAAAACAAAGA 1775
 QY 2274 AAAAAAGAAAGTGTGCGCCATTTAGCCGAGATCAATGAGCCCAAGCATATCAAGAA 2333
 Db 1776 AAAAAAGAAAGTGTGCGCCATTTAGCCGAGATCAATGAGCCCAAGCATATCAAGAA 1835
 QY 2334 TGTGTGCTGGAATTTTAAACCAATGATGATTTGACATGAGCAGCAAGATGCTTCAACA 2393
 Db 1836 TGTGTGCTGGAATTTTAAACCAATGATGATTTGACATGAGCAGCAAGATGCTTCAACA 1895
 QY 2394 CGAAGTTGAGCTGATGATGAAACAAGTGTGATGACAGAGTTCCTCCATTCACA 2453
 Db 1896 CGAAGTTGAGCTGATGATGAAACAAGTGTGATGACAGAGTTCCTCCATTCACA 1955
 QY 2454 GTGTGATGACCCCGCCGCGCAGTCACTTACATTAAGTCAAGAAATGCTGACCTCAATA 2513
 Db 1956 GTGTGATGACCCCGCCGCGCAGTCACTTACATTAAGTCAAGAAATGCTGACCTCAATA 2015
 QY 2514 AATTAAGCCCTCTCTGATGCTCTGATGATGATGATGATGATGATGATGATGATGAT 2573
 Db 2016 AATTAAGCCCTCTCTGATGCTCTGATGATGATGATGATGATGATGATGATGATGAT 2075
 QY 2574 AGGCAAAATGATATTTGAAATATTTCTTAACCCGAGCAGATGATTAAGAAATTTTAA 2633
 Db 2076 AGGCAAAATGATATTTGAAATATTTCTTAACCCGAGCAGATGATTAAGAAATTTTAA 2135
 QY 2634 AGGTCGCAAAACCACTTTGATGATGATGATGATGATGATGATGATGATGATGAT 2693
 Db 2136 AGGTCGCAAAACCACTTTGATGATGATGATGATGATGATGATGATGATGATGAT 2195
 QY 2694 CTAAAGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2753
 Db 2196 CTAAAGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2255
 QY 2754 TTTGAAGAGAGCTGGGAGAGTGGCCATTAAGGAGGAGAGTCTTTCAGACCACTCT 2813
 Db 2256 TTTGAAGAGAGCTGGGAGAGTGGCCATTAAGGAGGAGAGTCTTTCAGACCACTCT 2315
 QY 2814 TAGAGGCAATCAGCAGGCTCCATCAATCAAGGAGATGAGATGATGATGATGATGAT 2873
 Db 2316 TAGAGGCAATCAGCAGGCTCCATCAATCAAGGAGATGAGATGATGATGATGATGAT 2375
 QY 2874 CTGATTTAAGGAATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2933
 Db 2376 CTGATTTAAGGAATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2435
 QY 2934 TTTAAGGCTTTTATTTATGTTTCATGGAAGAACTGCTTATGAT 2981
 Db 2436 TTTAAGGCTTTTATTTATGTTTCATGGAAGAACTGCTTATGAT 2483

RESULT 3
 AY102701 2552 bp mRNA linear ROD 23-UN-2002
 LOCUS AY102701
 DEFINITION Mus musculus embryonic growth-associated protein EGAP mRNA,
 complete cds.
 ACCESSION AY102701
 VERSION AY102701.1 GI:21539895
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2552)
 AUTHORS Wenzlau, J.M. and Weisler-Evans, M.C.M.
 TITLE Mouse Embryonic Growth Associated Protein (EGAP)
 JOURNAL 2 (bases 1 to 2552)
 REFERENCE Wenzlau, J.M. and Weisler-Evans, M.C.M.
 AUTHORS Direct Submission
 TITLE Submitted (13-MAY-2002) Pediatrics, University of Colorado HSC,
 JOURNAL 4200 East Ninth Ave, Denver, CO 80262, USA
 FEATURES 1..2552
 SOURCE /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /dev_stage="embryo"
 55..2232
 /codon_start=1
 /product="embryonic growth-associated protein EGAP"
 /protein_id="AA052342.1"
 /db_xref="GI:2153896"

CDS
 /translation="WMKATVDDSDSGWELGVPEKXKSTSWVDITODFEDAGRELK
 /translating="WMKATVDDSDSGWELGVPEKXKSTSWVDITODFEDAGRELK
 /db_xref="GI:2153896"

BASE COUNT 740 a 542 c 619 g 651 t

Query Match 59.4%; Score 1838.4; DB 10; Length 2552;
 Best Local Similarity 84.4%; Pred. No. 0;
 Matches 2160; Conservative 0; Mismatches 341; Indels 57; Gaps 6;

1021 ATCTGTGACATTGCAAGGAAAAAGTAAATAAGCTGCTGTTTGAAGAGAGATTT 1080
 Db ATCTGACATTGCAAGGAAAAAGTAAATAAGCTGCTGTTTGAAGAGAGATTT 579
 520 ATCTGACATTGCAAGGAAAAAGTAAATAAGCTGCTGTTTGAAGAGAGATTT 579
 1081 CAGTCAATGACTTATGATTTAAATGGCTAAAGTGGAGAGATCTTGAAGTACAGC 1140
 Db CAGTCAATGACTTATGATTTAAATGGCTAAAGTGGAGAGATCTTGAAGTACAGC 639
 580 CAGTCAATGACTTATGATTTAAATGGCTAAAGTGGAGAGATCTTGAAGTACAGC 639
 1141 ATGCTTAAAGATGGAGAGATGACATGCAAGAGAGATTAAGAGTACTGAAAGTACAA 1200
 Db ATGCTTAAAGATGGAGAGATGACATGCAAGAGAGATTAAGAGTACTGAAAGTACAA 699
 640 ATGCTTAAAGATGGAGAGATGACATGCAAGAGAGATTAAGAGTACTGAAAGTACAA 699
 1201 GGAGAGAGAGATGCAAGAGATGCAAGAGATGCAAGAGATGCAAGAGATGCAAGAG 1260
 Db GGAGAGAGAGATGCAAGAGATGCAAGAGATGCAAGAGATGCAAGAGATGCAAGAG 759
 700 GGAGAGAGAGATGCAAGAGATGCAAGAGATGCAAGAGATGCAAGAGATGCAAGAG 759
 1261 AGAGTGAATTTACTGCTGCTTACAGAGAGATGCAAGAGATGCAAGAGATGCAAGAG 1320
 Db AGAGTGAATTTACTGCTGCTTACAGAGAGATGCAAGAGATGCAAGAGATGCAAGAG 819
 760 AGAGTGAATTTACTGCTGCTTACAGAGAGATGCAAGAGATGCAAGAGATGCAAGAG 819
 1321 AGAGTGAATTTACTGCTGCTTACAGAGAGATGCAAGAGATGCAAGAGATGCAAGAG 1380
 Db AGAGTGAATTTACTGCTGCTTACAGAGAGATGCAAGAGATGCAAGAGATGCAAGAG 879
 820 AGAGTGAATTTACTGCTGCTTACAGAGAGATGCAAGAGATGCAAGAGATGCAAGAG 879
 1381 CATATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1440
 Db CATATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 939
 880 CATATATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 939
 1441 ATTATGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGAT 1500
 Db ATTATGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGAT 999
 940 ATTATGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGAT 999
 1501 TATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
 Db TATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1059
 1000 TATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1059
 1561 ATTAATAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1620
 Db ATTAATAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1119
 1060 ATTAATAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1119
 1621 TGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1680
 Db TGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1179
 1120 TGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1179
 1681 CTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1740
 Db CTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1239
 1180 CTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1239
 1741 CTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1800
 Db CTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1297
 1240 CTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1297
 1801 ACCAGGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1860
 Db ACCAGGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1357
 1298 ACCAGGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1357
 1861 TTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1920
 Db TTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1417
 1358 TTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1417
 1921 AGAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1980
 Db AGAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1454
 1418 AGAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1454
 1981 TGAATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2040
 Db TGAATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1514
 1455 TGAATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1514
 2041 TTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2100
 Db TTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1574
 1515 TTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1574
 2101 TTATTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2160


```

Db      1575 CTTTGAGCTGGAGCTGACAGCATGCTAGTACTACTACATCTACTGCTGACTCTCCGA 1634
Qy      2161 ATTTCCTTACGAGCTGTTGATGTCACATGATGCGCGGATGCGCTCCAAATGGCAGA 2220
Db      1635 GTTCTGTATGACAGCTGATGTCACATGCTGAGCGCGCTGACGCGCTCTCAGATGGCAGA 1694
Qy      2221 GGAAGGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
Db      1695 GGAAGGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1754
Qy      2281 GAAAGTTCGCGCATTTGAGCGAGAGATCACAATGAGCGCAAGCATATCAGAACTGTGTGC 2340
Db      1755 GAAAGTTCGCGCATTTGAGCGAGAGATCACAATGAGCGCGCATATCAGAACTGTGTGC 1814
Qy      2341 TGAAGTTCGCGCATTTGAGCGAGAGATCACAATGAGCGCGCATATCAGAACTGTGTGC 2400
Db      1815 TGGGATGTTCAAGACGATGGTGGCTTTGACATGATGGCAAGTGGCAAGCCAAAGTT 1874
Qy      2401 TGAAGTTCGCGCATTTGAGCGAGAGATCACAATGAGCGCGCATATCAGAACTGTGTGC 2460
Db      1875 TGAAGTTCGCGCATTTGAGCGAGAGATCACAATGAGCGCGCATATCAGAACTGTGTGC 1934
Qy      2461 GACCGCGCGCGAGCTGACACTACTTACAGTCAAGAAATGCTGACCTCAATAAATATAG 2520
Db      1935 GACCGCGCGCGAGCTGACACTACTTACAGTCAAGAAATGCTGACCTCAATAAATATAG 1994
Qy      2521 CCTCTCTCTCTGAGTCTCTCTGAGCTGTATGTCGACGCTAGTATGACACTTTTCAACAGC 2580
Db      1995 CCTCTCTCTCTGAGTCTCTCTGAGCTGTATGTCGACGCTAGTATGACACTTTTCAACAGC 2054
Qy      2581 AATGATATGGAAGATATATCTCTAATCCGAGACCATGAGTAAATGAAATTTTAAAGGTTC 2640
Db      2055 AATGATATGGAAGATATATCTCTAATCCGAGACCATGAGTAAATGAAATTTTAAAGGTTC 2114
Qy      2641 CAAGCCCACTTTGTGTTATGAAAGTATTGTCGAGAGAGACACAAAGAGAAATCTTAAAGT 2700
Db      2115 CAAGCCCACTTTGTGTTATGAAAGTATTGTCGAGAGAGACACAAAGAGAAATCTTAAAGT 2174
Qy      2701 TCCCTCTGGAATTTGATTTCTCTGCTCAATAATATTTTCTGTTGTTGTTGTTGTTGTTGTT 2760
Db      2175 TCCCTCTGGAATTTGATTTCTCTGCTCAATAATATTTTCTGTTGTTGTTGTTGTTGTTGTT 2234
Qy      2761 GAGACTGGAGAGGTGGCCATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2820
Db      2235 GAGACTGGAGAGGTGGCCATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2269
Qy      2821 CACATCCACGAGCTCCACATCAGC--GGAAGTGAATGGAATTTCTGAGTAAACATCAT 2878
Db      2270 GACATCCACGAGCTCCACATCAGTGAAGAGAGTGAATCTTGAATTAACAGAGCTTGC 2329
Qy      2879 TATAAGGAATCTTTAGTTTACAGCCTTATATGACATGAATGAATGAATGAATGAATGAAT 2938
Db      2330 TACAGGAAGACTTTTACAGCCTTATATGACATGAATGAATGAATGAATGAATGAATGAAT 2389
Qy      2939 AGTGGTATATATGATGATGAGAGAGAACTGTCTTATGATGATGATGATGATGATGATGAT 2998
Db      2380 AGTGGTATATGATGATGATGAGAGAGAACTGTCTTATGATGATGATGATGATGATGATGAT 2445
Qy      2999 ATGCTTTTATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3058
Db      2446 ATGCTTTTATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2504
Qy      3059 GTTATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3096
Db      2505 GTTATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2542

```

```

ACCESSION AF272892
VERSION AF272892.1 GI:8926319
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 2543)
  Y.I.X.J., Li,X.F. and Yu,F.S.
  A novel epithelial wound-related gene is abundantly expressed in
  developing rat cornea and skin
  JOURNAL Curr. Eye Res. 20 (5), 430-440 (2000)
  MEDLINE 20314365
  PUBMED 10855038
REFERENCE
  2 (bases 1 to 2543)
  Li,X. and Yu,F.
  Direct Submission
  JOURNAL Submitted (26-MAY-2000) The Schepens Eye Research, 20 Staniford
  Street, Boston, MA 02114, USA
FEATURES
  source
    1..2543
    /organism="Rattus norvegicus"
    /mol_type="mRNA"
    /db_xref="taxon:10116"
    64..2241
    /codon_start=1
    /product="corneal wound healing related protein"
    /protein_id="AA81791.1"
    /db_xref="GI:8926320"
    /translation="MWMKAAVDDASGELNVPKEMKXSTSVVDITODFEDACRELK
    LGELHDKLFLGFEMSAIEMMDPMDAGMIGNVRYLNPEQAKDGTIKIDLSL
    PELGIMDFECCLITWDEGSLAQVTEFCYLHNPDIEDPAMKAPLGLIKIDIA
    RKNNKAAAFEBEDPSMTYGRKXNAGYTDLPVGMLEMDMDMORVRYSTPSRGE
    RDPVELEHQCCLAFSRVYKTRPVLLIATTKETSAVAQKLMQALBSAII
    TSLHNGIQONGTIGDHPIMNFEPLVNQRLPPTFPYAKIIRKEVNTFLRID
    RIKTICEVNNLENLHCILDFCFSESGSPCVISRLDTPLVNDKRVKGTLMQDMV
    KDAPSFVSPVLSKPCILYNNHQAQKIDISFVTCVPCFSLVQIHGNRRLRODKL
    GHILEEFAFLQDEAKVDALHTMLKQPOGQHLACGTWLVLSLRIMIOYLSGF
    DIELYSMEHYIYVLYSEPLVYAMKSTLSRADGSOAEERIMEEOQGRSKTKKK
    KVRPLREITWISQAYKQMGKMPYVAFPMDDGVRRPKFELDSQVRYEHRFAPFN
    SYMTPPPHYIOLFKMSDLSKSPPPQPELVYAS GFQOKMLLESTPNADREVS
    ILTKVKNVVKLLAGGKSKSKVPEEDSVHRTFVVKLV"
BASE COUNT 725 a 554 c 627 g 637 t
ORIGIN
Query Match 58.1%; Score 1799.2; DB 10; Length 2543;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 2193; Conservative 0; Mismatches 343; Indels 64; Gaps 7;
Qy 550 GGCATATGTTATGAAAGCTTCTGTAGATGATGACGATTCAGATGAGAGCTCAGATG 609
Db 58 GGCATATGTTATGAAAGCTTCTGTAGATGATGACGCTTCGAGATGAGAGCTCAGATG 117
Qy 610 CCAGAAAAATGAGAAAAAGCAATCAAACTGGGTGGAATTAACCAAGATTGAGAA 669
Db 118 CCGAAGAAAAATGAGAAAAAGCAATCAAACTGGGTGGAATTAACCAAGATTGAGAA 177
Qy 178 GCTTGTGAGAGAGTGAAGTGGAGAACTGTCAGATTAAGCTGTTGTGCTGTTGAA 237
Db 670 GCTTGTGAGAAATTAAGTGGAGAACTGTCAGATTAAGCTGTTGTGCTGTTGAA 729
Qy 730 GGCATGCTGCTATGAAATGATGATCCCAAGATGATGATGATGATGATGATGATGATGAT 789
Db 238 GGCATGCTGCTATGAAATGATGATCCCAAGATGATGATGATGATGATGATGATGATGAT 297
Qy 298 GCTTATGAAAAAGTCTCAATTTTGAACAAGCGCTCAAGATGAGACCAATTAATTA 357
Db 790 GTTATGAAAAAGTCTCAATTTTGAACAAGCTATCAAGATGAGACCAATTAATTA 849
Qy 298 GTTATGAAAAAGTCTCAATTTTGAACAAGCGCTCAAGATGAGACCAATTAATTA 357
Db 850 GACCTACCTTGCTGCACTGATAGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 909
Qy 358 GACCTACCTTGCTGCACTGATAGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 417

```

RESULT 4
 AF272892 2543 bp mRNA linear ROD 15-SEP-2000
 LOCUS Rattus norvegicus corneal wound healing related protein mRNA,
 DEFINITION complete cds.


```

QY 910 TGGTGAAGGCCATTCACCTGGCAGACAGATTTTACGCGCTTACATCATATCA 969
DB 418 TGGCTGGAAGGCCATTCCTGGCAGACAGATTTTACATGCTTACATCATATCA 477
QY 970 GACTTTATAGAGATCCTGCTATGAAGGCTTTTGTCTGCGAACTTTGAAAATCTGTGAC 1029
DB 478 GACTTTATAGAGATCCTGCTATGAAGGCTTTTGTCTGCGAACTTTGAAAATCTGTGAC 537
QY 1030 ATTGCAAGGAAAAGTAAATTAAGCTGCTTTTGAAGAGAGATTTTACAGTCAATG 1089
DB 538 ATTGCAAGGAAAAGTCAACAAAGCTGCTTTTGAAGAGAGATTTTCAAGTCAATG 597
QY 1090 ACTTATGATTTTAAAATGCTTAAACAGTGTGACAGATCTTTGAGTTACAGGCAATGCTAAA 1149
DB 598 ACATATGATTTTAAAATGCTTAAACAGTGTGACAGATCTTTGAGTTACAGGCAATGCTAAA 657
QY 1150 GATGTGAGGATGACATGCAAGAAAGATTAAGATCGAAGTGCAGAGAGAGAA 1209
DB 658 GATGTGAGGATGACATGCAAGAGAGATTAAGATCGAAGTGCAGAGAGAGAG 717
QY 1210 AGAGATCCAGAGATTAAGATTAAGATCGAAGTGCAGAGATTAAGATCGAAGAG 1269
DB 718 AGAGATCCAGAGATTAAGATTAAGATCGAAGTGCAGAGATTAAGATCGAAGAG 777
QY 1270 TTTACTGCTGTGTACTGACAGTGTATAGCTTTACTAAGAAAAGAGAGAGTGTGT 1329
DB 778 TTACACCGAGTACTGCTGACAGTGTCTATAGCTTTACTAAGAAAAGAGAGAGTGTGT 837
QY 1330 GCAGAGCTCAAAAATTTGATGTTTCAAGAGAGATCTTTTCCGCTTCAATATCA 1389
DB 838 GCAGAGCTCAAGAACTGATGTTTCAAGAGAGAGATCTTTTCCGCTTCAATATCA 897
QY 1390 TTGATCATGAGATCCAGAGCTCAAGATTAAGATCGAAGTGCAGAGATTAAGAT 1449
DB 898 TTGATCATGAGATCCAGAGCTCAAGATTAAGATCGAAGTGCAGAGATTAAGAT 957
QY 1450 GGTTTTGAAGCCCTTGTGAACAGAGCTACTTCCACTACCTTCCCTGATATGCAAA 1509
DB 958 GGTTTTGAAGCCCTTGTGAACAGAGATTAAGATCGAAGTGCAGAGATTAAGAT 1017
QY 1510 ATATTTAAAGGAGAAATGTTGAACTTTTGAAGATTAAGATTAAGATTAAGAT 1569
DB 1018 ATATTTAAAGGAGAAATGTTGAACTTTTGAAGATTAAGATTAAGATTAAGAT 1077
QY 1570 GTCTGAGGTTGTGAATTAACAAATTTACATGATCTGATTTTCTGTGATTT 1629
DB 1078 GTCTGAGGTTGTGAATTAACAAATTTACATGATCTGATTTTCTGTGATTT 1137
QY 1630 AGTGAACAGTCAAGATGTTTCTTGAAGATCTGTTGAAACCACTTCTGTGTGAT 1689
DB 1138 AGTGAACAGTCAAGATGTTTCTTGAAGATCTGTTGAAACCACTTCTGTGTGAT 1197
QY 1690 AAAAAAAGTCTTTTGAAGATCTGTTGAAAGATGTTGAAAGATGTTGAAAGAT 1749
DB 1198 AAAAAAAGTCTTTTGAAGATCTGTTGAAAGATGTTGAAAGATGTTGAAAGAT 1257
QY 1750 TTGTGAGATCTCTGAGTGTCTTCCCAAGTGTACTATATATATATATATATAT 1809
DB 1258 TTGTGAG--TCTCTGAGTGTCTCTCCCAAGTGTACTATATATATATATATATAT 1315
QY 1810 AGGATGTTTGAAGTCTTGTGTTTGAAGTGTGTTGAGGATTTGTTGTTTGA 1869
DB 1316 AGGATGTTTGAAGTCTTGTGTTTGAAGTGTGTTGAGGATTTGTTGTTTGA 1375
QY 1870 TCCATGACATTAACAGAGCTGCAACAGAGATTAAGTGTGTTTGAAGATTTG 1929
DB 1376 TCCATGACATTAACAGAGCTGCAACAGAGATTAAGTGTGTTTGAAGATTTG 1435
QY 1930 CCACCTTGAAGATGTTTATGACATTTTATATAGGAGAGAGATGTTGAGC 1989
DB 1436 CCACCTTGAAGATGTTTATGACATTTTATATAGGAGAGATGTTGAGC 1472

```

```

QY 1990 GCTTCAACCATGCTGTGTAAGAAACAGAAACCCCAAGCAATTTGCGCTTTAGTAC 2049
DB 1473 GCTTCAACCATGCTGTGTAAGAAACAGAAACCCCAAGCAATTTGCGCTTTAGTAC 1532
QY 2050 CTGGGCTCTTTACCATTAACCTTCCGATTATGATACAGTACCTTTCTAGTGTGAAAT 2109
DB 1533 TTGGGCTCTTTACCATTAACCTTCCGATTATGATACAGTACCTTTCTAGTGTGAAAT 1592
QY 2110 GGAATCTTACAGTATGACAGTACTATTAATATATATATATATATATATATATATAT 2169
DB 1593 GGAATCTTACAGTATGACAGTACTATTAATATATATATATATATATATATATATAT 1652
QY 2170 GCGATGTTGATGTCACATTAAGTGTGCGGATGAGCTCTCAAAATGCGAGAGAAAGAT 2229
DB 1653 GCGATGTTGATGTCACATTAAGTGTGCGGATGAGCTCTCAAAATGCGAGAGAGAT 1712
QY 2230 AATGGAAGAGAGAGAAAGGCGTGTAGTAAATTAAGAAATTAAGAAATTAAGAAAT 2289
DB 1713 AATGGAAGAGAGAGAAAGGCGTGTAGTAAATTAAGAAATTAAGAAATTAAGAAAT 1772
QY 2290 CCGATTTGAGGAGAGATCAAGATGAGCCAGATATCAAGATGTTGCTGAGATTT 2349
DB 1773 TCCATTTGAGGAGAGATCAAGATGAGCCAGATATCAAGATGTTGCTGAGATTT 1832
QY 2350 TAAAAATGTTAGTATTTGACATGAGAGCAAGATTAAGATTAAGATTTGAGCTTGA 2409
DB 1833 TAAAAATGTTAGTATTTGACATGAGAGCAAGATTAAGATTAAGATTTGAGCTTGA 1892
QY 2410 TAGTGAACAAATTTGATGTAACAGAGTGTGCTTCAATTAAGATGATGAGAGAGAG 2469
DB 1893 TAGTGAACAAATTTGATGTAACAGAGTGTGCTTCAATTAAGATGATGAGAGAGAG 1952
QY 2470 GCGATGCTACTTACTTCAAGTTCAGAAATGTTGAGCTTCAATTAAGATTTGAGCTTCC 2529
DB 1953 GCGATGCTACTTACTTCAAGTTCAGAAATGTTGAGCTTCAATTAAGATTTGAGCTTCC 2012
QY 2530 TCAGTCTGTAACATGTTGTTGAGAGCTGATTAAGATTTGAGAGAGAGAGAGATTT 2589
DB 2013 TCAGTCTGTAACATGTTGTTGAGAGCTGATTAAGATTTGAGAGAGAGAGATTT 2072
QY 2590 GGAATAATTTCTTAACCCGAGCATGAGTTTATAGATTTTAAAGTTTGCACAAACCAA 2649
DB 2073 GGAATAATTTCTTAACCCGAGCATGAGTTTATAGATTTTAAAGTTTGCACAAACCAA 2132
QY 2650 CTGTTGTTTGAAGTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2709
DB 2133 CTGTTGTTTGAAGTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2192
QY 2710 ATTGATTTCTGCTGATTAATATTTTCTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 2769
DB 2193 ATTGATTTCTGCTGATTAATATTTTCTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 2252
QY 2770 GAGGTGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2829
DB 2253 GAGGTGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2283
QY 2830 AGGCTCCACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2886
DB 2284 AGGCTCCACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2343
QY 2887 -ATATTTTATGTTGACAGCTTATATGACATGATGAAGAGAGAGAGAGAGAGAG 2945
DB 2344 -ATATTTTATGTTGACAGCTTATATGACATGATGAAGAGAGAGAGAGAGAGAG 2402
QY 2946 TATATGTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3005
DB 2403 TATATGTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2458
QY 3006 TATTAAGATTTATTAACAAATCAATTTTATATGAGATTAAGAGAGAGAGAGAGAG 3065
DB 2459 TATTAAGATTTATTAACAAATCAATTTTATATGAGATTAAGAGAGAGAGAGAGAG 2517
QY 3066 AAGGTTAATAATTTCTTGAACAAAA 3091

```

```

Db      2518  AAGCTTAATAATTCTTGACAAAA 2543
|||||
RESULT 5
LOCUS   AK025266
DEFINITION Homo sapiens CDNA: FLJ21613 f1s, clone COL07381.
ACCESSION AK025266
VERSION   AK025266.1 GI:10437744
KEYWORDS  oligo capping: f1s (full insert sequence).
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS  Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Otagaki,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
JOURNAL  Unpublished
TITLE     2 (bases 1 to 1839)
COMMENT   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
DIRECT SUBMISSION
TITLE     Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
JOURNAL   University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@life.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5' & 3' end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source    1..1839
            /location=Qualifiers
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="COL07381"
            /tissue_type="colon"
            /clone_id="COL"
            /note="Cloning vector pME18sFL3"
            134..1714
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="BAB15097.1"
            /db_xref="GI:10437745"
            /translation="VVMKASVDDSDSGELSMPEKMEKSTNTNWVDITDPEACRELK
            LGEILHDKLPGLFEMASAIEMDPKMDAGMIGNORVULFPEALIKGTIKIDITLT
            PELGIMTGCOCCLITMLEGSLQVATVCTYIHNPDIEDBPAKARLGIKICDIP
            REKNKAARFEEDPQSMITGPKFANSTDLRVTGMLDVAEDDMORKSTRSGOEE
            RDPELEHQRCLAVFSVKTRTALVLIIVLPTFKETSAVAEOLIVQADLLSAIH
            NSLHGIIQAONDITTKGDIIVGWGPELVNQSLPTPEPRVAKI1KREMNVPFRLID
            RIKTVCSEVNLNTNHCIDLPFCSEFSQPCVLSLSLOTFLVNNKQVFTHLMQDMW
            KDAIRFSEVPSPLSKRYCLYNNHQAQCDIDSFVTHCVPPFCSLIQIHGNFARORDKI
            GHILEEFAITDDEAKKVDAAALHTMLLNKQEPORHLAWIGTWLVHNLAIMIYLLSGS
            ELKYSKHEILHITIVSL"
BASE COUNT 565 a 336 c 437 g 501 t
ORIGIN
Query Match 57.1%; Score 1766.6; DB 9; Length 1839;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1828; Conservative 0; Mismatches 9; Indels 26; Gaps 3;
423 ATACGATCGCGTCACGCGTCGCGGCTGAGAGAGGAGGAGGCGGCGGCGG 482
|||||
Db      1 ATACGATCGCGTCACGCGTCGCGGCTGAGAGAGGAGGAGGCGGCGGCGG 60
483 CGAGCGGCGTCGTTATTTCCGTGGTCCGGACAGTCGCGTGCAGCGGAGTACCAACGAG 542

```

Db	61	CGAGGCGGCGCTGTTATTTCCGTGTCGGGACAGTGCGTGCGCGCGGGCTGCACAGCG	120
Qy	543	AGAGATAGCATATATGGTTATGAAAGCTCTGTAGATGATGAGATTCACAGATGGAGCT	602
Db	121	AGAGATGAGCATATATGGTTATGAAAGCTCTGTAGATGATGAGATTCACAGATGGAGCT	180
Qy	603	CAGTATCCGAAAAAATGAGAAAAAGCATATACAACTGGGTGAGCATTAACCAAGATT	662
Db	181	CAGTATCCGAAAAAATGAGAAAAAGCATATACAACTGGGTGAGCATTAACCAAGATT	240
Qy	663	TGAAAGAGCTTGTGCGAATTAAGTTGGAGAACTACTCATGATTAAGTATTGGCT	722
Db	241	TGAAAGAGCTTGTGCGAATTAAGTTGGAGAACTACTCATGATTAAGTATTGGCT	300
Qy	723	TTTGAAGCCATGTCTGCTATTGAAATGATGATCCCAAGATGGATGCTGCATATGG	782
Db	301	TTTGAAGCCATGTCTGCTATTGAAATGATGATCCCAAGATGGATGCTGCATATGG	360
Qy	783	AAACCAAGTTAATCGAAAAATTTCTCAATTTTGAACAGCTATCAAGATGGCATATTA	842
Db	361	AAACCAAGTTAATCGAAAAATTTCTCAATTTTGAACAGCTATCAAGATGGCATATTA	420
Qy	843	AATTAAGATCTACCTTGCCTGACACTGATAGGAGATTATGATACATAGTTTGTGCTTT	902
Db	421	AATTAAGATCTACCTTGCCTGACACTGATAGGAGATTATGATACATAGTTTGTGCTTT	480
Qy	903	GATPACGTGGTTGAAGAGCATTCACCTGGACAGACAGATTTTAACGTCCTTACATTA	962
Db	481	GATPACGTGGTTGAAGAGCATTCACCTGGACAGACAGATTTTAACGTCCTTACATTA	540
Qy	963	TAATCCAGACTTTATAGAGATCCTGCTATGAAGGCTTTTGTCTGGGAACTTTGAAAT	1022
Db	541	TAATCCAGACTTTATAGAGATCCTGCTATGAAGGCTTTTGTCTGGGAACTTTGAAAT	600
Qy	1023	CTGTGACATTTGCAAGGAAAAAGTAAAGCTGCTTTTGAAGAAGAAATTTTCA	1082
Db	601	CTGTGACATTTGCAAGGAAAAAGTAAAGCTGCTTTTGAAGAAGAAATTTTCA	660
Qy	1083	GTCAATCACTTATGATTTAAAAATGCGTAAACAGTGTACAGATCTTCAGATTACAGCAT	1142
Db	661	GTCAATCACTTATGATTTAAAAATGCGTAAACAGTGTACAGATCTTCAGATTACAGCAT	720
Qy	1143	GCTTAAAGATGTGAGAGATGACATGCAAAAGAGTAAAGTAAAGTCTCGAATCGACAAAG	1202
Db	721	GCTTAAAGATGTGAGAGATGACATGCAAAAGAGTAAAGTAAAGTCTCGAATCGACAAAG	780
Qy	1203	AGAAAGAAAGATCCAGAGTTGACTAGAACACCAACAAATGTTTAGCAGTATTACAGAG	1262
Db	781	AGAAAGAAAGATCCAGAGTTGACTAGAACACCAACAAATGTTTAGCAGTATTACAGAG	840
Qy	1263	AGTGAATTTTACTGCTGTTACTGACAGAGCTTATGAGCTTTACTTAAAGAAAGAGACAG	1322
Db	841	AGTGAATTTTACTGCTGTTACTGACAGAGCTTATGAGCTTTACTTAAAGAAAGAGACAG	900
Qy	1323	TGCTGTTGCAAGAGCTCAAAAATTTGATGTTTCAAGACAGATCTTCTTCTGCGATTCA	1382
Db	901	TGCTGTTGCAAGAGCTCAAAAATTTGATGTTTCAAGACAGATCTTCTTCTGCGATTCA	960
Qy	1383	TAATTCATTCGATATGAGATCCAGAGCCCAAGATGATCTACAAAGAGATCAATCCAT	1442
Db	961	TAATTCATTCGATATGAGATCCAGAGCCCAAGATGATCTACAAAGAGATCAATCCAT	1020
Qy	1443	TATGATGGGTTTGGAAACCCCTTGTGAACCAAGAGCTATTCACATTAACCTTCCCTGATA	1502
Db	1021	TATGATGGGTTTGGAAACCCCTTGTGAACCAAGAGCTATTCACATTAACCTTCCCTGATA	1080
Qy	1503	TGCAAAAAATATTTAAAAAGGAGAAATGAGAACTATTTTGAACAAGATTAAATAGTAGAAT	1562
Db	1081	TGCAAAAAATATTTAAAAAGGAGAAATGAGAACTATTTTGAACAAGATTAAATAGTAGAAT	1140
Qy	1563	AAAAAAGTCTGTGAGGTTGTGAATTTAACAAATTTACATGTATCCTGSAATTTTTTCTG	1622

Db 1141 AAAACGTCTGAGAGGTGTGAATTAAACAATTACATTGATCTGATTTTCTG 1200

Qy 1623 TGAATTTAGTGAACAGTACACATGTTCTTTCAAGATCTCTGTTACAAACATTTCT 1682

Db 1201 TGAATTTAGTGAACAGTACACATGTTCTTTCAAGATCTCTGTTACAAACATTTCT 1260

Qy 1683 GGTGATTAACAAAGAGTCTTTGAACTCATCTGACAGACATGTTGAAAGATGACAT 1742

Db 1261 GGTGATTAACAAAGAGTCTTTGAACTCATCTGACAGACATGTTGAAAGATGACAT 1320

Qy 1743 TCGGCTTTTGTGAGTCTCTCGAGTCTTTTCCCAAGTCTTACCTATATATATAC 1802

Db 1321 TCGGCTTTTGTGAGTCTCTCGAGTCTTTTCCCAAGTCTTACCTATATATATAC 1378

Qy 1803 CAGGCTAAGAGCTGTATGATCTCTTGTATCACTGTTGGGCACTTGTGATCTT 1862

Db 1379 CAGGCTAAGAGCTGTATGATCTCTTGTATCACTGTTGGGCACTTGTGATCTT 1438

Qy 1863 ATTACATCCATGATGACATTAACAGAGGCTGACAGAGATTAAGCTTGTATATCTT 1922

Db 1439 ATTACATCCATGATGACATTAACAGAGGCTGACAGAGATTAAGCTTGTATATCTT 1498

Qy 1923 GAATTTGCCACCTTGAGAGTGTATGACATTTATTTAATAGAGAGAGTTG 1982

Db 1499 GAATTTGCCACCTTGAGAGTGTATGACATTTATTTAATAGAGAGAGTTG 1535

Qy 1983 ATGACAGCGCTTACACACATGCTGTTGAAACAGAAACCCAAAGCAATTTGGCTGTT 2042

Db 1536 ATGACAGCGCTTACACACATGCTGTTGAAACAGAAACCCAAAGCAATTTGGCTGTT 1595

Qy 2043 TAGTACCTGGGCTCTTACCATTAACCTTGCATTAATGATACAGTCTTCTTAAGTCT 2102

Db 1596 TAGTACCTGGGCTCTTACCATTAACCTTGCATTAATGATACAGTCTTCTTAAGTCT 1655

Qy 2103 TTGAATTTGACACTTACATGATGACAGAG-TACTATTACATATATGATCTCTGAA 2161

Db 1656 TTGAATTTGACACTTACATGATGACAGAG-TACTATTACATATATGATCTCTGAA 1715

Qy 2162 TTGCTTTAGCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2221

Db 1716 TTGCTTTAGCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1775

Qy 2222 GAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2281

Db 1776 GAAAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1835

Qy 2282 AAA 2284

Db 1836 AAA 1838

RESULT 6

BC053286 2467 bp mRNA linear VRT 09-JUN-2003

LOCUS BC053286

DEFINITION Dantio rerio cDNA clone MGC:64157 IMAGE:6797204, complete cds.

ACCESSION BC053286

VERSION BC053286.1 GI:31418975

KEYWORDS MGC.

SOURCE Dantio rerio (zebrafish)

ORGANISM Dantio rerio

REFERENCE

1 (bases 1 to 2467)

Strausberg, R., Feingold, E.A., Grouse, L.H., Berge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Scheffer, C.F., Bharti, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stachleiron, M., Soares, M.B., Bonaldo, M.F., Casavant, T.D., Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loguettano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wotley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, S., Krzyzanski, M.I., Skalska, U., Smal, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marz, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE 22388257

PUBMED 12472932

2 (bases 1 to 2467)

Strausberg, R.

Direct Submission

Submitted (02-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Leonard I. Zon, M.D.

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lln.gov>

Series: IRAC Plate: 117 Row: n Column: 11

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, similarity but not identity to protein.

FEATURES

source

location/qualifiers

1..2467

/organism="Dantio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

/clone="MGC:64157 IMAGE:6797204"

/rname="zebrafish"

/clone_lib="NCI CGAP_ZKID1"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6.1"

58..2232

/codon_start=1

/product="unknown (protein for MGC:64157)"

/protein_id="AA53286.1"

/db_xref="GI:31418976"

/translation="MVKSSVBEDEGGGICTEPKKRNANWDTVDFEGACKELK GELHDKLFGFEMASAIEMWDPRMDGNGVNRVNLFDQALVKEALRVKDLSTP ELGIMDTCCCLITWLEHSLAQVTCLVYHNPDIIDPALKAFALGILKICDIR EKVKVAAFEEDFOAMTYGFKMANVYTDLVHGMKDVDEDLQKRVKSTRSGEOR DPEVELDHQOCALFSRVTFRLLLSALISFTKKTSAVSEADQKQKADSLPAAVA TIQVGIQNDITTKDHPIMWGEPELVNORLPTTPRYAKIIRKEMVAFSLIER IKSVCEIINTLHSLIDFCESSEOSPCLSLSITOTTELLINKKVFGRHLMODMK DALRCFVSPVLSKSNHNAKDVIJSVPHCTRPFSLSIQIHANARORDKXG HILBERPTDDEAKVDALHGLMKLEPDRQHLACGIVLTHNLKIMVQYLSGE LELVSKHEIYIYTWSEFLVAMWSTLSADSSQVAREILLEQLKVRSSKSKKK KARPLSEITMSQAYQMCAWYKTMALMDRKVRKQPELDSQVRYEHRAPAPS VVTPPPHYIOFEMSLDKKYNPPPSADLYMAASKHFOQAKLLELVTSPPDAEVARI LKVAKPIVVMKLAGGKKETKALPEFDSAKYPIYKVL"

BASE COUNT

720 a 574 c 600 g 573 t

Query Match 40.98; Score 1264.8; DB 5; Length 2467;

Best Local Similarity 74.3%; Pred. No. 8.3e-276;
Matches 1648; Conservative 0; Mismatches 542; Indels 28; Gaps 3;

543 AGAAGTATGATATGTTATGAAAGCTCTGTAGATGATGACGATTGAGATGGAGCT 602
Db AAAAGCTGTCTCATGTGTATGAAATGATCTGTGAGAGAGAGAGAGAGAGATGGGCT 104
45 AAAAGCTGTCTCATGTGTATGAAATGATCTGTGAGAGAGAGAGAGAGATGGGCT 104
QY CAGTATCCAGAAAAATGAGAAAAAGCAATACAACTGGGTGACATTACCAAGATT 662
Db TGGATCCAGAGAAAGT---GAGAAACAATGCCAAGTGGGTGAGATGCCAGAGATT 161
QY TGAAGAACTGTGAGAAATTAAGTTGGAGAACTACTTCATGATTAAGTATTGGTCT 722
663 TGAAGAACTGTGAGAAATTAAGTTGGAGAACTACTTCATGATTAAGTATTGGTCT 722
Db TAAAGGCTGTGAAAGAACTCAACTGGAGCTGTCCATGACAAAGCTATTGGGCT 221
162 TAAAGGCTGTGAAAGAACTCAACTGGAGCTGTCCATGACAAAGCTATTGGGCT 221
QY TTTTGAAGCATGTCTGTATTTGAATGATGATCCCAAGATGATGCTGGCATTTGG 782
723 TTTTGAAGCATGTCTGTATTTGAATGATGATCCCAAGATGATGCTGGCATTTGG 782
Db GTTGAAGCATGTCTGTATTTGAATGATGATCCCAAGATGATGCTGGCATTTGG 281
222 GTTGAAGCATGTCTGTATTTGAATGATGATCCCAAGATGATGCTGGCATTTGG 281
QY AAACCAAGTAAATGAAAAAGTTCTCAATTTTGAACAAAGCTATGAAGAGGCACTTAA 842
783 AAACCAAGTAAATGAAAAAGTTCTCAATTTTGAACAAAGCTATGAAGAGGCACTTAA 842
Db AAATCAAGTAAACCGTAAAGTCTCAACTTTGACCAAGCTGTTAAGAGTAAAGCCATAG 341
282 AAATCAAGTAAACCGTAAAGTCTCAACTTTGACCAAGCTGTTAAGAGTAAAGCCATAG 341
QY AATTAAAGATCTCACTGTGCTGATGATGATGAGGATTAATGATACATGTTTGTCTTT 902
843 AATTAAAGATCTCACTGTGCTGATGATGATGAGGATTAATGATACATGTTTGTCTTT 902
Db GGTGAAGATCTCACTGTGCTGATGATGATGAGGATTAATGATACATGTTTGTCTTT 401
342 GGTGAAGATCTCACTGTGCTGATGATGATGAGGATTAATGATACATGTTTGTCTTT 401
QY GATTAAGTGTGTAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 962
903 GATTAAGTGTGTAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 962
Db GATTAAGTGTGTAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 461
402 GATTAAGTGTGTAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 461
QY TAAATCCAGCTTTATAGAAAGCTCTGCTGATGAGAGGCTTTGCTGGGATCTTGAAT 1022
963 TAAATCCAGCTTTATAGAAAGCTCTGCTGATGAGAGGCTTTGCTGGGATCTTGAAT 1022
Db TAAATCCAGCTTTATAGAAAGCTCTGCTGATGAGAGGCTTTGCTGGGATCTTGAAT 521
462 TAAATCCAGCTTTATAGAAAGCTCTGCTGATGAGAGGCTTTGCTGGGATCTTGAAT 521
QY CTGTGATCTTGAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 1082
1023 CTGTGATCTTGAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 1082
Db CTGTGATCTTGAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 581
522 CTGTGATCTTGAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 581
QY GTCAATGATCTTGAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 1142
1083 GTCAATGATCTTGAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 1142
Db GTCAATGATCTTGAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 641
582 GTCAATGATCTTGAAGAGGCACTTCACTGAGCAAGAGTATTAAGTGTCTTACATTCA 641
QY GCTTAAAGATGTGAGAGTATGATGAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 1202
1143 GCTTAAAGATGTGAGAGTATGATGAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 1202
Db GCTTAAAGATGTGAGAGTATGATGAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 701
642 GCTTAAAGATGTGAGAGTATGATGAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 701
QY AGAAGAAAGAGTATGAGAGTATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 1262
1203 AGAAGAAAGAGTATGAGAGTATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 1262
Db TGAACAGAGAGTATGAGAGTATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 761
702 TGAACAGAGAGTATGAGAGTATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 761
QY AGTAAATTTAATCTGTGTCTTACTGACAGTCTTAAAGCTTTTAAAGAAAGAGTAAAG 1322
1363 AGTAAATTTAATCTGTGTCTTACTGACAGTCTTAAAGCTTTTAAAGAAAGAGTAAAG 1322
Db GGTAAATTTAATCTGTGTCTTACTGACAGTCTTAAAGCTTTTAAAGAAAGAGTAAAG 821
762 GGTAAATTTAATCTGTGTCTTACTGACAGTCTTAAAGCTTTTAAAGAAAGAGTAAAG 821
QY TGTGTTGAGAGTATGAGAGTATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 1382
1223 TGTGTTGAGAGTATGAGAGTATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 1382
Db TGTGTTGAGAGTATGAGAGTATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 881
822 TGTGTTGAGAGTATGAGAGTATGAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 881
QY TAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442
1383 TAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442
Db TGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 941
882 TGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 941
QY TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1502
1443 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1502
Db CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001
942 CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001
QY TGAAGAAATTAATTAAG 1562
1503 TGAAGAAATTAATTAAG 1562
Db TGCACAAATCATCAAG 1061
1002 TGCACAAATCATCAAG 1061
QY AAAAAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622
1563 AAAAAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1622

Db TAAAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1121
1062 TAAAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1121
QY TGAATTTAGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1682
1623 TGAATTTAGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1682
Db TGAATTTAGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1181
1122 TGAATTTAGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1181
QY GGTGATTAACAAAAGGCTTTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1742
1683 GGTGATTAACAAAAGGCTTTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1742
Db GATGACAAACAAAGGCTTTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1241
1182 GATGACAAACAAAGGCTTTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1241
QY TGTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1802
1743 TGTGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1802
Db GCGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1299
1242 GCGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1299
QY CAGGCTAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1862
1803 CAGGCTAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1862
Db CAGGCTAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
1300 CAGGCTAAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
QY ATTGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1922
1863 ATTGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1922
Db ATTCGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1419
1360 ATTCGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1419
QY GAATTTGCCACCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1982
1923 GAATTTGCCACCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1982
Db GAGTTGCCACCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1456
1420 GAGTTGCCACCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1456
QY ATGCAGGCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2042
1983 ATGCAGGCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2042
Db ACCGCTCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1516
1457 ACCGCTCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1516
QY TAGTACCTGAGTCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2102
2043 TAGTACCTGAGTCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2102
Db TAGTACCTGAGTCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1576
1517 TAGTACCTGAGTCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1576
QY TTGAATTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2162
2103 TTGAATTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2162
Db TTGAATTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1636
1577 TTGAATTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1636
QY TCCTTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2222
2163 TCCTTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2222
Db TCCTTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1636
1637 TCCTTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1636
QY AAAGGATTAATGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2282
2223 AAAGGATTAATGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2282
Db AAAGGATTAATGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1756
1697 AAAGGATTAATGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1756
QY AAGTTCAGGCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2342
2283 AAGTTCAGGCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2342
Db AAGTTCAGGCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1816
1757 AAGTTCAGGCTTGCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1816
QY GAATTTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2402
2343 GAATTTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2402
Db GAATTTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1876
1817 GAATTTTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1876
QY AGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2462
2403 AGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2462
Db AGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1936
1877 AGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1936
QY CCGCGCGCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2522
2463 CCGCGCGCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2522
Db CCGCGCGCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1996
1937 CCGCGCGCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1996
QY CTCCTCTCAGTCTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2582
2523 CTCCTCTCAGTCTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2582
Db CTCCTCTCAGTCTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
1997 CTCCTCTCAGTCTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
QY TGAATTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2642
2583 TGAATTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2642
Db TGAATTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2116
2057 TGAATTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2116
QY AACCCAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2702
2643 AACCCAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2702
Db AACCCAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2176
2117 AACCCAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2176

AUTHORS
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project

TITLE
Unpublished

JOURNAL
2 (bases 1 to 1796)

REFERENCE
Isogai, T., Otsuki, T. and Sugiyama, T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source
Location/Qualifiers
1..1796
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2NE2005317"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuron after the differentiation of NT2 neuronal precursor cells."

BASE COUNT
490 a 365 c 463 g 478 t

ORIGIN
Query Match 28.6%; Score 885.4; DB 9; Length 1796;
Best Local Similarity 99.9%; Pred. No. 9e-190;
Matches 886; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 549 AGGCATATGTTATGAAGCTTCTAGATGATGACGATTCGAGTGGAGCTCGTAT 608
DB 345 AGGCATATGTTATGAAGCTTCTAGATGATGACGATTCGAGTGGAGCTCGTAT 404
QY 609 GCCAGAAAAATGAGAAAAAGCAATCAAACTGGGAGCATACCAGATTGTAAGA 668
DB 405 GCCAGAAAAATGAGAAAAAGCAATCAAACTGGGAGCATACCAGATTGTAAGA 464
QY 669 AGCTTGTGGAATTAAGTTGGAGAACTACTTCAATGATAGCTATTGGTCTTTTGA 728
DB 465 AGCTTGTGGAATTAAGTTGGAGAACTACTTCAATGATAGCTATTGGTCTTTTGA 524
QY 729 AGGCATGCTGATTAATGAATGATGATCCCAAGATGATGCTGGCATGTTGGAACCA 788
DB 525 AGGCATGCTGATTAATGAATGATGATCCCAAGATGATGCTGGCATGTTGGAACCA 584
QY 789 AGTTATCGAAAAAGTTCTCAATTTTGAACAAGCTATCAAGATGCGACTTTAAATTA 848
DB 585 AGTTATCGAAAAAGTTCTCAATTTTGAACAAGCTATCAAGATGCGACTTTAAATTA 644
QY 849 AGATTCACCTGCTGCTAGATGATAGGATTAATGATACATGTTTGTGCTGTTGTAAC 908
DB 645 AGATTCACCTGCTGCTAGATGATAGGATTAATGATACATGTTTGTGCTGTTGTAAC 704
QY 909 GTGGTTAGAGGCGCATTCAGTGCACAGACGATTTTACGCTTTTACATTCAATATCC 968
DB 705 GTGGTTAGAGGCGCATTCAGTGCACAGACGATTTTACGCTTTTACATTCAATATCC 764
QY 969 AGACTTATAGAAAGATCTGCTATGAAGGCTTTGTCTCTGGAACTGTGAATAATCTGTA 1028

DB 765 AGACTTATAGAAAGATCTGCTATGAAGGCTTTGTCTCTGGAACTGTGAATAATCTGTA 824
QY 1029 CATTGCAAGGAAAAAGTAATAAAGCTGCTGTTTGAAGAGGAGATTTTCAGTCAAT 1088
DB 825 CATTGCAAGGAAAAAGTAATAAAGCTGCTGTTTGAAGAGGAGATTTTCAGTCAAT 884
QY 1089 GACTTATGATTTAAATGGCTTAACAGTGTGACAGATCTTTCAGTTACAGGATGCTTAA 1148
DB 885 GACTTATGATTTAAATGGCTTAACAGTGTGACAGATCTTTCAGTTACAGGATGCTTAA 944
QY 1149 AGATGTGAGAGATGACATGCAAGAGAGTAAGAGTACTGAAAGTGCACAGAGAGA 1208
DB 945 AGATGTGAGAGATGACATGCAAGAGAGTAAGAGTACTGAAAGTGCACAGAGAGA 1004
QY 1209 AAGAGATCAGAGATTGAACTAGAACACCAAAATGTTTAGCAATTTACAGAGAGTGA 1268
DB 1005 AAGAGATCAGAGATTGAACTAGAACACCAAAATGTTTAGCAATTTACAGAGAGTGA 1064
QY 1269 ATTACTGCTGTGTTACTGACAGAGTCTTATAGCTTTACTAGAAAGACCAAGTCTGT 1328
DB 1065 ATTACTGCTGTGTTACTGACAGAGTCTTATAGCTTTACTAGAAAGACCAAGTCTGT 1124
QY 1329 TGCAAGAGCTCAAAAATTGATGTTCAAGACGACAGATCTTTCTGCAATTCATATTC 1388
DB 1125 TGCAAGAGCTCAAAAATTGATGTTCAAGACGACAGATCTTTCTGCAATTCATATTC 1184
QY 1389 ATTGCATGATGCATCCAGGCCAGAAATGATACACAAAGAGATC 1435
DB 1185 ATTGCATGATGCATCCAGGCCAGAAATGATACACAAAGAGATC 1231

RESULT 9
AX664135 719 bp DNA linear PAT 22-MAR-2003
LOCUS AX664135
DEFINITION Sequence 7 from Patent WO02057304.
ACCESSION AX664135
VERSION AX664135.1 GI:29164145
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1 Panzer, S.R., Lincoln, S.E., Altus, C.M., Dufour, G.E., Jackson, J.L., Jones, A.L., Dam, T.C., Liu, T.F., Harris, B., Flores, V., Daiflo, A., Marwaha, R., Chen, A.J., Chang, S.C., Gerstin, E.H., Peralta, C.H., David, M.H. and Lewis, S.A.
Secretory molecules
Patent: WO 02057304-A 7 25-JUL-2002;
JOURNAL
Incyte Genomics, Inc. (US)
FEATURES
source
Location/Qualifiers
1..719
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: LI:462889.1:2001JAN12"
BASE COUNT
223 a 113 c 188 g 195 t

ORIGIN
Query Match 23.2%; Score 717.4; DB 6; Length 719;
Best Local Similarity 99.9%; Pred. No. 9.9e-152;
Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 474 GCGCGCGCGCGAGGCGGCGTATTTCGCGGTCGCGACAGTGCAGGCGCGCGGCT 533
DB 1 GCGCGCGCGCGAGGCGGCGTATTTCGCGGTCGCGACAGTGCAGGCGCGGCT 60
QY 534 GACCAAGGAGAAAGTGAAGCATTAATGTTATGAAAAGCTTCTAGATGATGACATTCAG 593
DB 61 GACCAAGGAGAAAGTGAAGCATTAATGTTATGAAAAGCTTCTAGATGATGACATTCAG 120
QY 594 ATGGAGCTCAGTATGCGCAAGAAAAATGAGAAAAAGCAATCAAACTGGGTGAGCATTAC 653

```

Db      121 ATGGGAGCTCAGTATCCAGAAAAATGAGAAAAACATACAAACTGGGTGACATTTC 180
QY      654 CCAAGATTTTGAAGAGCTTGTGAGAAATTAAGTTGGAGAACTAATTCATGATTAAGCT 713
Db      181 CCAAGATTTTGAAGAGCTTGTGAGAAATTAAGTTGGAGAACTAATTCATGATTAAGCT 240
QY      714 ATTGGCTTTTGAAGAGCTTGTGAGAAATTAAGTTGGAGAACTAATTCATGATTAAGCT 773
Db      241 ATTGGCTTTTGAAGAGCTTGTGAGAAATTAAGTTGGAGAACTAATTCATGATTAAGCT 300
QY      774 CATGATTTGAAACCAAGTTAATGAAAGTTCTCAATTTTGAACAAGCTATCAGAGATG 833
Db      301 CATGATTTGAAACCAAGTTAATGAAAGTTCTCAATTTTGAACAAGCTATCAGAGATG 360
QY      834 CATATTTAAATTAAGATCTGCTTGTGAGAAATTAAGTTGGAGAACTAATTCATGATTAAGCT 893
Db      361 CATATTTAAATTAAGATCTGCTTGTGAGAAATTAAGTTGGAGAACTAATTCATGATTAAGCT 420
QY      894 TTGCTGTTTGAATGAGCTGTTGAGAAAGCTTCACTGAGACAGAGATTTTACGTGCT 953
Db      421 TTGCTGTTTGAATGAGCTGTTGAGAAAGCTTCACTGAGACAGAGATTTTACGTGCT 480
QY      954 TTACATTTCAATCCAGACTTTAATGAAATCTGCTATGAAAGCTTTTGTCTGGAGAT 1013
Db      481 TTACATTTCAATCCAGACTTTAATGAAATCTGCTATGAAAGCTTTTGTCTGGAGAT 540
QY      1014 CTGAAATCTGATGATTCAGAGGAGAAAGTAAATTAAGCTGCTTTTGAAGAGA 1073
Db      541 CTGAAATCTGATGATTCAGAGGAGAAAGTAAATTAAGCTGCTTTTGAAGAGA 600
QY      1074 AGATTTTCACTGATTCATGATTTAATGAAATCTGCTATGAAAGCTTTTGTCTGGAGAT 1133
Db      601 AGATTTTCACTGATTCATGATTTAATGAAATCTGCTATGAAAGCTTTTGTCTGGAGAT 660
QY      1134 TACAGCATGCTAAAGATGTGAGAGATGATGACATGCAAGAGAGTAAAGACTGCGAA 1192
Db      661 TACAGCATGCTAAAGATGTGAGAGATGATGACATGCAAGAGAGTAAAGACTGCGAA 719

RESULT 10
AC102010 147876 bp DNA linear HTG 21-MAR-2003
LOCUS Mus musculus clone RP24-531W24, WORKING DRAFT SEQUENCE, 7 unordered
DEFINITION pieces
AC102010 GI:29135693
VERSION HTG; HTGS PHASE1; HTGS DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 147876)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP24-531W24
Unpublished
2 (bases 1 to 147876)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Chopel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gaidyna,S.,
Ginde,S., Goid,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,
Lamaze,R., Lander,E., Lander,T., Lechoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margis,N., Matthews,C.,
McCarthy,M., McEwan,T., McKernan,K., McPherson,R., Melidim,J.,
Menus,L., Minova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Ratta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 147876)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Chopel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Deatellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gaidyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Lander,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McEwan,T., Menus,L., Minova,T., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Ratta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tefaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 21, 2003 this sequence version replaced gs:28412092.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 531_M_24
Center clone name: U17873
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 14609 bases at least Q40
Consensus quality: 14686 bases at least Q30
Consensus quality: 14704 bases at least Q20
Insert size: 15700; agarose-fp
Insert size: 147276; sum-of-coverage
Quality coverage: 9.9 in Q20 bases; agarose-fp
Quality coverage: 10.6 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 84958: contig of 84958 bp in length
* 84959 85058: gap of 100 bp
* 85059 87411: contig of 2353 bp in length
* 87412 87511: gap of 100 bp
* 87512 88933: contig of 2422 bp in length
* 88934 90023: gap of 100 bp
* 90024 94077: contig of 4043 bp in length
* 94077 94176: gap of 100 bp

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: November 25, 2003, 03:13:51 | Search time 106 Seconds
(without alignments)
10683.169 Million cell updates/sec

Title: US-10-001-857-42

Perfect score: 5439
Sequence: 1 tttcttcagcaactcccaag.....attcttcgacaaaaaaaaa 3096

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 667484 seqs, 182883703 residues

Total number of hits satisfying chosen parameters: 1334968

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=kh
-Q=/cgn2/USFTO.spool/US10001857/runat.24112003.162856.9623/app_query.fasta_1.3271
-DB=published.Applications_AA -QFT=fastan -SUFFIX=n2p.iapb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS-human40.cdi -List=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=200000000 -USER=US10001857 @CGN 1.1.149 @runat.24112003.162856.9623
-NCPU=6 -ICPU=3 -NO MAP -LARGESOURCE -NEG SCORES=0 -WAIT -DEPBLCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	3796	69.8	733	14	US-10-001-857-145	Sequence 145, App
2	165	3.0	523	12	US-10-017-161-1982	Sequence 1982, Ap
3	146.5	2.7	574	12	US-10-168-097A-76	Sequence 76, Appl
4	146.5	2.7	574	12	US-10-239-431A-38	Sequence 38, Appl
5	145	2.6	559	12	US-09-858-155A-2	Sequence 2, Appl1
6	145	2.6	559	12	US-10-168-097A-26	Sequence 26, Appl
7	145	2.6	559	12	US-10-239-431A-33	Sequence 33, Appl
8	145	2.6	559	12	US-10-116-370-2	Sequence 2, Appl1
9	144	2.6	1134	12	US-10-001-873-50	Sequence 50, Appl
10	142.5	2.6	800	12	US-10-029-386-32198	Sequence 32198, A
11	142	2.6	721	14	US-10-086-464-5	Sequence 5, Appl1
12	141.5	2.6	5877	15	US-10-142-515-11	Sequence 11, Appl
13	141.5	2.6	5935	9	US-10-243-243A-8	Sequence 8, Appl1
14	140	2.5	802	9	US-09-823-240-2	Sequence 2, Appl1
15	139.5	2.5	1243	15	US-10-196-935A-4	Sequence 4, Appl1
16	138.5	2.5	641	11	US-09-959-987-2	Sequence 2, Appl1
17	138.5	2.5	647	9	US-09-925-297-1002	Sequence 1002, Ap
18	138.5	2.5	647	11	US-09-925-299-1002	Sequence 1002, Ap
19	136.5	2.5	647	14	US-10-086-464-2	Sequence 2, Appl1
20	136.5	2.5	647	14	US-10-086-464-4	Sequence 4, Appl1
21	134.5	2.4	520	12	US-10-168-097A-36	Sequence 36, Appl
22	134.5	2.4	520	12	US-10-239-431A-34	Sequence 34, Appl
23	134	2.4	351	12	US-10-239-431A-22	Sequence 22, Appl
24	134	2.4	542	12	US-10-239-431A-10	Sequence 10, Appl
25	133.5	2.4	507	14	US-10-078-547-24	Sequence 24, Appl
26	132.5	2.4	206	15	US-10-156-761-14830	Sequence 14830, A
27	132.5	2.4	507	15	US-10-156-761-10907	Sequence 10907, A
28	132	2.4	503	14	US-10-078-547-2	Sequence 2, Appl1
29	131	2.4	528	12	US-09-840-746-20	Sequence 20, Appl
30	131	2.4	538	10	US-09-976-740-43	Sequence 43, Appl
31	131	2.4	538	14	US-10-023-529-43	Sequence 43, Appl
32	131	2.4	538	14	US-10-023-523-43	Sequence 43, Appl
33	130	2.4	455	12	US-10-032-585-7399	Sequence 7399, Ap
34	129.5	2.4	501	12	US-10-168-097A-46	Sequence 46, Appl
35	129.5	2.4	377	15	US-10-239-431A-35	Sequence 35, Appl
36	129	2.3	377	15	US-10-149-819-7	Sequence 7, Appl1
37	129	2.3	505	12	US-10-168-097A-56	Sequence 56, Appl
38	129	2.3	505	12	US-10-239-431A-36	Sequence 36, Appl
39	126.5	2.3	405	12	US-10-017-161-2256	Sequence 2256, Ap
40	126.5	2.3	1046	15	US-10-156-761-10088	Sequence 10088, A
41	126.5	2.3	416	15	US-10-043-487-282	Sequence 282, App
42	126	2.3	505	12	US-10-168-097A-12	Sequence 12, Appl
43	126	2.3	505	12	US-10-239-431A-32	Sequence 32, Appl
44	125	2.3	505	12	US-09-976-740-47	Sequence 47, Appl
45	125	2.3	550	10	US-09-976-740-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-10-001-857-145
Sequence 145, Application US/10001857
Publication No. US20020183500A1
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Herve
APPLICANT: Chen, Sei-Yu
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and p
FILE REFERENCE: DEK-0273
CURRENT APPLICATION NUMBER: US/10/001,857
PRIOR FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/252,054
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.1
SEQ ID NO 145
LENGTH: 733
TYPE: PRT
ORGANISM: Homo sapien
US-10-001-857-145

Alignment Scores:

Pred. No.:	5e-294	Length:	733
Score:	3796.00	Matches:	732
Percent Similarity:	99.59%	Conservative:	0
Best Local Similarity:	99.59%	Mismatches:	1
Query Match:	69.79%	Indels:	3
	14	Gaps:	0

US-10-001-857-42 (1-3096) x US-10-001-857-145 (1-733)

```

QY 556 ATGTTTGAAGAGCTTCTGATGATGACGATTCAGAGTGGGAGCTCAGATGGCCGAA 615
DB 1 MetValIleuLysAlaSerValAspAspAspSerGlyTrpGluLeuSerMetProGlu 20
QY 616 AAAATGGAGAAACCAATACAACTGGGTGGACATTACCAAGATTTTGAAGAGCTTGT 675
DB 21 LysMetGluLysSerAsnThrAsnTrpValAspIleThrGlnAspPheGluGluAlaCys 40
QY 676 CGAAGATTAAAGTTGGAGAACTACTCATGATTAAGCTATTGGCTTTTGAAGCATG 735
DB 41 ArgGluLeuLysLeuGlyGluLeuLeuHisAspLysLeuPheGlyLeuPheGluAlaMet 60
QY 736 TCTGCTATTGAAATGATGATGCCAAGATGGATGCTGGCATGATTGGAAACCAATTAT 755
DB 61 SerAlaIleGluMetMetAspProLysMetAspPheAspAlaGlyMetIleGlyAsnGlnValAsn 80
QY 796 CGAAAGTTCTCAATTTTGAACCAAGCTATCAAGATGGCATATTAAATTAAGATCTC 855
DB 81 ArgLysValIleuAsnPheGluGlnAlaIleLysAspGlyThrIleLysIleLysAspLeu 100
QY 856 ACCTGCTGAAGATGATAGGATTAATGATACATGTTTGTCTGTTGATTAAGCTGTTA 915
DB 101 ThrLeuProGluLeuIleGlyIleMetAspThrCysPheCysLeuIleThrTrpLeu 120
QY 916 GAAGCCATTCTACCTGGCACAGACAGTATTACGTGCTTTACATTATCCAGACTTT 975
DB 121 GluGlyHisSerLeuAlaGlnThrValPheThrCysLeuTyrlleHisAsnProAspPhe 140
QY 976 ATAGAAGATCCTGCTATGAAGGCTTTTCTGGAATCTTGAAGATCTGACATTCGA 1035
DB 141 IleGluAspProAlaMetLysAlaPheAlaLeuGlyIleLeuLysIleCysAspIleAla 160
QY 1036 AGGGAAAAAGTAAATAGCTGCTGTTTGAAGAGAAATTTTCAGTCAATGCTTAT 1095
DB 161 ArgGluLysValAsnLysAlaAlaValPheGluGluGluAspPheGlnSerMetThrTyr 180
QY 1096 GGATTTAAATGGCTAAACAGTGTGACAGATCTTCAGATTACAGGATGCTAAAAGATGTG 1155
DB 181 GlyPheLysMetAlaAsnSerValThrAspLeuAspValThrGlyMetLeuLysAspVal 200
QY 1156 GAGGATGACATGCAAAAGAGTAAAGTACTGGAAGTGCACAAAGAAAGAAAGAT 1215
DB 201 GluAspAspMetGlnArgValLysSerThrArgSerArgGlnGlyGluGlnArgAsp 220
QY 1216 CCAGAAGTTGAAGTGAACACCAACATGTTTAGAGATTCAGAGAGTGAAGTAAATTA 1275
DB 221 ProGluValGluLeuGluHisIleGlnIleCysLeuAlaValPheSerArgValLysPheThr 240
QY 1276 CGTGTGTACTGACAGTCTTATAGCTTTTACTTAAGAAAGAGACAGTGTGTTGCAAA 1335
DB 241 ArgValIleuLeuThrValIleuIleAlaPheThrLysLysGluThrSerAlaValAlaGlu 260
QY 1336 GCTCAAAATGATGGTTCAGACAGACAGATCTTCTTTCGCAATTCATATTCATGAT 1395
DB 261 AlaGlnLysLeuMetValGlnAlaAlaAspLeuLeuSerAlaIleHisAsnSerLeuHis 280
QY 1366 CATGGCATCCAGCCCAAGATGATATCTAAGAAAGAGATCATCCATATGATGGGTTT 1455
DB 281 HisGlyIleGlnAlaGlnAsnAspThrThrLysGlyAspHisPheGlyMetGlyPhe 300
QY 1456 GAACCCCTTGTAACACAGAGCTACTTCACCTACTCTCCGTGATTAAGAAATTAAT 1515
DB 301 GluProLeuValAsnGlnArgLeuLeuProPheThrPheProArgTyrAlaLysIleIle 320

```

```

QY 1516 AAAAGGAGAAATGATGAACTATTTTGCAAGATTATAGATAGATAAAACCTGCTGT 1575
DB 321 LysArgGluGluMetValAsnTrpPheAlaArgLeuIleAspArgIleLysThrValCys 340
QY 1576 GAGCTTGTAAATTTAAACAAATTTACATTGTATCCCTGGATTTTTTCTGCAATTTGTA 1635
DB 341 GluValValAsnLeuThrAsnLeuHisCysIleLeuAspPhePheCysGluPheSerGlu 360
QY 1636 CAGTCAACATGTGTCTTCAAGATCTCTGTATCAAAACCACTTCTCGTGGATTAACAA 1695
DB 361 GlnSerProCysValLeuSerArgSerLeuLeuGlnThrThrPheLeuValAspAsnLys 380
QY 1696 AAGTCTTTGAACTCATCTTCAGACAGATGGTGAAGAAAGTGCACCTTGGCTTTTGTG 1755
DB 381 LysValPheGlyThrHisLeuMetGlnAspMetValLysAspAlaLeuArgSerPheVal 400
QY 1756 AGATCCTCCGAGTGGCTTTCCCAAGTGGTACTATATATATATATATATATATATATAT 1815
DB 401 -AspProPro-ValLeuSerProLysCysTyrLeuTyrlAsnAsnHisGlnAlaLysAspC 420
QY 1816 GTATGCACTCCTTTGTACTCACTGTGTGGCCATCTGTAGCTTATTCAGATCCATG 1875
DB 420 YelIleAspSerPheValThrHisCysValArgProPheCysSerLeuIleGlnIleHisG 440
QY 1876 GACATACAGGGCTGACAGAGATTAAGCTTGTCATATCTTGAAGAAATTTGCCACCT 1935
DB 440 LysHisAsnArgAlaArgGlnArgAspLysLeuGlyHisIleLeuGluGluPheAlaThrL 460
QY 1936 TGCAGATGAGTTTATGACATTTATTTTATATAGCA -GAGAAGTTGATGACAGCCTTC 1994
DB 460 euGlnAspGluPheMetThrPheTyrPheAsnArgAlaGluLysValAspAlaAlaLeuH 480
QY 1995 ACACCATGCTGTGGAACAGAACCCCAAGCAACATTTGGCTGTTTAAAGTAACTGGG 2054
DB 480 IsthThrLeuLeuLysGlnGluProGlnArgGlnHisLeuAlaCysLeuGlyThrTrpY 500
QY 2055 TCCTTACCATTAACCTTCGCATTTATATATAGTACATCTTGAAGTGGCTTTGAATTGAA 2114
DB 500 alleuTyrlHisAsnLeuArgIleMetIleGlnTyrlLeuLeuSerGlyPheGluLeuGlu 520
QY 2115 TCTACAGTATGACAGAGTCTATTATATATTTGATATCTCTCTGAATTCCTTTAGCAT 2174
DB 520 euTyrlSerMetHisGlnTyrlTyrlTyrlIleTyrlTrpTyrlLeuSerGluPheLeuTyrlAla 540
QY 2175 GATTGATGTCACATGATGATCGGCGATGGCTCTCAAAATGCGACAGAGAAAGATTAATG 2234
DB 540 TrpLeuMetSerThrLeuSerArgAlaAspGlySerGlnMetAlaGlnAlaArgIleMetG 560
QY 2235 AAGAGCAGCAAGAAAGCCGTAGTATGTAATAAAACAAAGAAAGAAAGTTCGCCCAT 2294
DB 560 IuGluGlnGlnLysGlyArgSerSerLysThrLysLysLysLysValArgProL 580
QY 2295 TGAGCCGAGAGATCACATGAGCAGCATATGACAGACATGATGCTGCTGAATTTAAAA 2354
DB 580 euSerArgGluIleThrMetSerGlnAlaTyrlGlnHisMetCysAlaGlyMetPheLys 600
QY 2355 CCATGTGATGATTTGACATGAGCGCAAAAGTACGTAAACCGAAGTTTGAAGCTTGATAG 2414
DB 600 hrMetValAlaPheAspMetAspGlyLysValArgLysProLysPheGluLeuAspSerG 620
QY 2415 AACCAAGTTGGTATGAACACAGGTTTGCTCCATTCATCAACAGTGTGAGACCCGCCAG 2474
DB 620 IuGlnValAlaGlyTrpGluHisArgPheAlaProPheAsnSerValMetThrProPro 640
QY 2475 TGCATTAATTACAGTTCAAGGAAATGTCTGACCTCAATAAATATAGCCCTCTCTCAGT 2534
DB 640 alHisTyrlLeuGlnPheLysGluMetSerAspLeuAsnLysTyrlSerProProGlnHis 660
QY 2535 CTCCTGAATGTATGTGGAGCAGCTAGTAAGCACTTCAACAGGCAAAATGATTTGANA 2594
DB 660 eTrpGluLeuTyrlValAlaAlaSerLysHisIsthPheGlnGlnAlaLysMetIleLeuGlu 680

```

QY 2595 ATATTCCTAACCCGACCATGAGTTAAATTTAAAGTTGCCAAACCACTTTG 2654
 DB 680 snileProAsnProkaphisgiValAsnArgileLeuLyValAlaLeysProAsnPhv 700
 QY 2655 TGGTATTAAGTTATTTGGCAGAGACACAAAGAACTTAAAGTTCTCTGAAATTTG 2714
 DB 700 alValMeLysLeuValagLygLyshLysLysGlsSerLysValProProGluPhv 720
 QY 2715 ATTCTCTGCTCAATATATTTCTCTGTGGAACCTTGT 2755
 DB 720 sphenSerAlaHisLysTyrPheProValValLysLeuVal 733
 RESULT 2
 US-10-017-161-1982
 ; Sequence 1982, Application US/10017161
 ; Publication No. US2003014368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SUMA, MAKIKO
 ; APPLICANT: ASAI, KIYOSHI
 ; APPLICANT: AKIYAMA, YUTAKA
 ; APPLICANT: ABURAIANI, HIROYUKI
 ; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
 ; FILE REFERENCE: 084335/0152
 ; CURRENT APPLICATION NUMBER: US/10/017,161
 ; PRIOR FILING DATE: 2002-12-18
 ; PRIOR APPLICATION NUMBER: JP 2001/246789
 ; NUMBER OF SEQ ID NOS: 2430
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1982
 ; LENGTH: 523
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: MOD_RES
 ; LOCATION: (259)..(265)
 ; OTHER INFORMATION: Variable amino acid
 ; FEATURE:
 ; NAME/KEY: MOD_RES
 ; LOCATION: (373)..(376)
 ; OTHER INFORMATION: Variable amino acid
 US-10-017-161-1982
 Alignment Scores:
 Pred. No.: 0.000233 Length: 523
 Score: 165.00 Matches: 74
 Percent Similarity: 38.32% Conservative: 67
 Best Local Similarity: 20.11% Mismatches: 159
 Query Match: 3.00% Indels: 69
 DB: 12 Gaps: 8
 US-10-001-857-42 (1-3096) x US-10-017-161-1982 (1-523)
 QY 1337 GCTTTCGACAGCAGCTGCTCTCTTTAGTAAAGCTATAGACCTGTCAGTAACA 1278
 DB 3 AlathrAlathrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 14
 QY 1277 CGAGTAAATTCCTACTGCTGTAATCTGTAACATGTTGTTGTTAGTCACTTCT 1218
 DB 15 Thr 34
 QY 1217 GGATCTCTTCTCTCTCTGTCGATGAGTACTTACTCTTCTTCTGATGATCC 1158
 DB 35 Thr 54
 QY 1157 TCCACATCTTTTACAGCTCTGTAACCTGGAAGATCTGTCACATGTTAGCAATTTAAAT 1098
 DB 55 Thr 72
 QY 1097 CCATAGATCATGACTGAAATCTTCCTCAAAAACAGAGAGCTTATTTACTTTTCC 1038
 DB 73 -----Thr 83

QY 1037 CTTCGATGTCAGAGATTTTCAGATTTCCAGACAAAGCCCTTCATAGCAGATCTCT 978
 DB 84 IleIleIleThrAlaThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 103
 QY 977 ATTAAGCTGATTAATGATGTAAGACAGTAATTAAGTCTGTCTGCTGCAATGAGCTT 918
 DB 104 IleThr-----IleThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 112
 QY 917 TCTAACCCAGCTATCAACAGCAAAACATGTATCCATATTCCTATTCAGTCAAGCA 858
 DB 113 ThrIleThr 132
 QY 857 GTGAGATCTTTAATTTTATAGTCCATCTGATGCTTGTTCATAAATTGAGACTTTT 798
 DB 133 Thr 152
 QY 797 CGATTAAGCTGTTCCATGATGCGACGATCCATCTGAGATCCATTCATTCATAGCA 738
 DB 153 ThrAlaThr 172
 QY 737 GACATGCTTCAAAAAGACCAATAGCTTATCATGAGTAGTTCTCCCACTTTAATCT 678
 DB 173 AsnAsnAsnAsnAsnAsnAsnAsn-----AsnHisSerAsnAsnTyrIleThr 189
 QY 677 CGACAGCTTCTCAAAATCTTGGAATGTCACCCAGTTGTATTTCTTCAT 618
 DB 190 Thr 206
 QY 617 TTTTCGCACTGAGCTCCATCTGATCCATCATCATCATCATCATCATCATCATCATCAT 558
 DB 207 Thr 226
 QY 557 ATTATGCTTACTT-CTCCGCTGTCACCCGCGCCGCGCCGCGCCGCGCCGCGCCGCG 499
 DB 226 IshIshGlnIshIshThrProProProProProProProProProProProProProProProPro 246
 QY 498 TAAAGCAG-----491
 DB 246 hrThr 266
 QY 490 -----CGGCTGCG 454
 DB 266 roPro 286
 QY 453 AGCCCGACCGGACGCGACGATCGATCGCGGAGCGCGACCGCGCGCGCGCGCGCGCG 394
 DB 286 roPro 306
 QY 393 CTTTACTCCGGGAGCTTCCGCGCTCCCGCGCACACCCCTCCGCGCACCGCGCTCGT 334
 DB 306 roPro 325
 QY 333 TCGGCGATGCGCGCGCT 314
 DB 326 -----ProPro 327
 RESULT 3
 US-10-168-097A-76
 ; Sequence 76, Application US/10168097A
 ; Publication No. US20030166245A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CENTRE NATIONAL OF THE RECHERCHE SCIENTIFIQUE
 ; APPLICANT: INSTITUT CURIE
 ; FILE REFERENCE: IFB99MASP
 ; CURRENT APPLICATION NUMBER: US/10/168,097A
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 76
 ; LENGTH: 574
 ; TYPE: PRT
 ; ORGANISM: Schizosaccharomyces pombe

Percent Similarity: 40.52% Conservative: 14
Best Local Similarity: 31.37% Mismatches: 61
Query Match: 2.64% Indels: 30
DB: 10 Gaps: 8

US-10-001-857-42 (1-3096) x US-09-858-155A-2 (1-559)

```
QY 559 CCATTATGCTA-----CTTCTCCGTGATCAGCCGCGCGCAGCACT 515
DB 294 ProThrcysileserseralathrglyleuilegluasnargproglinsersproalathr 313
QY 514 GTCGCGACCGGAAATAAGACGCGCCCTCGGCGCGCCCGCCCTCCCTCTCAGCC 455
DB 314 GlyArgThrProValPheValSerProThrProProProProProProProProPro 333
QY 454 CAGCCGACCGGAGCGGTGCAGCATGCTATGCCGGAGCGCGACCC-----GCC 401
DB 334 AlaLeuSerThrserserleuArgAlaSerMetThrsThrProProProProValPro 353
QY 400 CCCTTCCTTACTCCGCGGAGACTCG---CGTCCCGCGCACACCCCT----- 353
DB 354 ProProProProProProProProProProProProProProProProProProPro 373
QY 352 -----CCTCGCACCG-----CCTCGCTTC 332
DB 374 ProLeuGlnIleAlaProGlyValLeuHisProAlaProProProIleAlaProProLeu 393
QY 331 GCGCGATGCGCGCGCTCTCTCTCGGAGATTAGACAGTCGCGAGCGGAGATT 272
DB 394 ValGlnProSerProPro-----ValAlaArgAlaAlaProValCys-----GluThr 409
QY 271 ATCTTTTACCCCTCTGTGCGAGTCTGAGTGGCGGTTCGCGGAAAGAGCGGAG 212
DB 410 ValProValHisProLeuProGlnGlyValGlnGlyLeuProProProProProPro 429
QY 211 CCGGAGTCTCAGAGCGCGCGCTGTGCGTGCCTCT 173
DB 430 ProProLeu-----ProProProGlyIleArgProSer 440
```

RESULT 6

US-10-168-097A-26
; Sequence 26, Application US/10168097A
; Publication No. US20030166245A1
; GENERAL INFORMATION:
; APPLICANT: CENTRE NATIONAL OF THE RECHERCHE SCIENTIFIQUE
; APPLICANT: INSTITUT CORIE
; TITLE OF INVENTION: WASP FAMILY PROTEIN FRAGMENTS, AND USES THEREOF
; FILE REFERENCE: IFB99WASP
; CURRENT APPLICATION NUMBER: US/10/168,097A
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-168-097A-26

Alignment Scores:
Pred. No.: 0.00942 Length: 559
Score: 145.00 Matches: 48
Percent Similarity: 40.52% Conservative: 14
Best Local Similarity: 31.37% Mismatches: 61
Query Match: 2.64% Indels: 30
DB: 12 Gaps: 8

US-10-001-857-42 (1-3096) x US-10-168-097A-26 (1-559)

```
QY 559 CCATTATGCTA-----CTTCTCCGTGATCAGCCGCGCGCAGCACT 515
DB 294 ProThrcysileserseralathrglyleuilegluasnargproglinsersproalathr 313
QY 514 GTCGCGACCGGAAATAAGACGCGCGCTGTGCGGCGCGCGCGCGCTCCCTCTCAGCC 455
```

```
DB 314 GlyArgThrProValPheValSerProThrProProProProProProProProSer 333
QY 454 CAGCCGACCGGAGCGGTGCAGCATGCTATGCCCGGAGCGCGACCC-----GCC 401
DB 334 AlaLeuSerThrserserleuArgAlaSerMetThrsThrProProProProProPro 353
QY 400 CCCTTCCTTACTCCGCGGAGACTCG---CGTCCCGCGCACACCCCT----- 353
DB 354 ProProProProProProProProProProProProProProProProProProPro 373
QY 352 -----CCTCGCACCG-----CCTCGCTTC 332
DB 374 ProLeuGlnIleAlaProGlyValLeuHisProAlaProProProIleAlaProProLeu 393
QY 331 GCGCGATGCGCGCGCTCTCTCTCGGAGATTAGACAGTCGCGAGCGGAGATT 272
DB 394 ValGlnProSerProPro-----ValAlaArgAlaAlaProValCys-----GluThr 409
QY 271 ATCTTTTACCCCTCTGTGCGAGTCTGAGTGGCGGTTCGCGGAAAGAGCGGAG 212
DB 410 ValProValHisProLeuProGlnGlyValGlnGlyLeuProProProProProPro 429
QY 211 CCGGAGTCTCAGAGCGCGCGCTGTGCGTGCCTCT 173
DB 430 ProProLeu-----ProProProGlyIleArgProSer 440
```

RESULT 7

US-10-239-431A-33
; Sequence 33, Application US/10239431A
; Publication No. US20030170726A1
; GENERAL INFORMATION:
; APPLICANT: FRADELIZE, JULIE
; APPLICANT: FRIEDERICH, EVELYNE
; APPLICANT: COLSTEYN, ROY M.
; APPLICANT: LOUARD, DANIEL
; APPLICANT: NOIREAUX, VINCENT
; APPLICANT: SYRES, CECILE
; TITLE OF INVENTION: PEPTIDE SEQUENCES COMPRISING ONE OR MORE UNITS BINDING
; TITLE OF INVENTION: TO PROTEINS OF THE ENA/WASP FAMILY, AND THEIR USES
; FILE REFERENCE: 0508-1032
; CURRENT APPLICATION NUMBER: US/10/239, 431A
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/FR01/00843
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: FR 00/03637
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-431A-33

Alignment Scores:
Pred. No.: 0.00942 Length: 559
Score: 145.00 Matches: 48
Percent Similarity: 40.52% Conservative: 14
Best Local Similarity: 31.37% Mismatches: 61
Query Match: 2.64% Indels: 30
DB: 12 Gaps: 8

US-10-001-857-42 (1-3096) x US-10-239-431A-33 (1-559)

```
QY 559 CCATTATGCTA-----CTTCTCCGTGATCAGCCGCGCGCAGCACT 515
DB 294 ProThrcysileserseralathrglyleuilegluasnargproglinsersproalathr 313
QY 514 GTCGCGACCGGAAATAAGACGCGCGCTGTGCGGCGCGCGCGCGCTCCCTCTCAGCC 455
DB 314 GlyArgThrProValPheValSerProThrProProProProProProProProSer 333
```

QY 454 CAGCCCGACGGAGGCTGACATGCTATGCCGGGACGGAGCCG-----GCC 401
 DB 334 AAlaLeuSerThrsSerLeuAlaSerMetThrsProProProProValPro 353
 QY 400 CCCTTCCTTACTCCGGGAGCTTG---CGCTCCCGGCGACACCCCT----- 353
 DB 354 ProProProProProProAlaThrAlaLeuGlnAlaProAlaValProProProAla 373
 QY 352 -----CCCTCGCGACCG-----CCTCGGCT 332
 DB 374 ProLeuGlnLeuAlaProGlyValLeuHisProAlaProProProAlaProProLeu 393
 QY 331 GCGCGATGGCGCGCGCTTCTTCTCGGAGATTAGACGATCGGAGCGGAATT 272
 DB 394 ValGlnProSerProPro-----ValAlaArgAlaAlaProValCys-----GluThr 409
 QY 271 ATCTTTTTCACCTCTGTGCGAGTCTCTGAGTGGGGGCTTCCGGAAGAGGGGAG 212
 DB 410 ValProValHisProLeuProGlnGlyGluValGlnGlyLeuProProProProPro 429
 QY 211 CCGGAGTCTCAGAGCCCGCGCGCTGCGCTGCCCTCT 173
 DB 430 ProProLeu-----ProProProGlyLeuArgProSer 440

RESULT 8

US-10-116-370-2
 ; Sequence 2, Application US/10116370
 ; Publication No. US20030190709A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; TITLE OF INVENTION: PABLO, A POLYPEPTIDE THAT INTERACTS WITH BCL-XL, AND
 ; FILE REFERENCE: AM100012-D2
 ; CURRENT APPLICATION NUMBER: US/10/116, 370
 ; PRIORITY FILING DATE: 2002-04-04
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 559
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-116-370-2

Alignment Scores:

Pred. No.: 0.00942 Length: 559
 Score: 145.00 Matches: 48
 Percent Similarity: 40.52% Conservative: 14
 Best Local Similarity: 31.37% Mismatches: 61
 Query Match: 2.64% Indels: 30
 DB: 12 Gaps: 8

US-10-001-857-42 (1-3096) x US-10-116-370-2 (1-559)

QY 559 CCATTATGCGCTA-----CTTCTCCCGGAGTACACCGCGCGCGCGAGCACT 515
 DB 294 ProThrCysIleSerSerAlaThrGlyLeuIleGlnAmArgProGlnSerProAlaThr 313
 QY 514 GTCCGACCAACGAATAACAGCGCGCTCGCGCGCGCGCGCGCTCCCTCTGAGCC 455
 DB 314 GlyArgThrProValPheValSerProThrProProProProProProProProLeuPro 333
 QY 454 CAGCCGACCGGAGGCTGACGATGCGCGGAGCGGCGGACCGC-----GCC 401
 DB 334 AAlaLeuSerThrsSerLeuAlaSerMetThrsProProProProValPro 353
 QY 400 CCCTTCCTTACTCCGGGAGCTTG---CGCTCCCGGCGACACCCCT----- 353
 DB 354 ProProProProProProAlaThrAlaLeuGlnAlaProAlaValProProProAla 373
 QY 352 -----CCCTCGCGACCG-----CCTCGGCT 332
 DB 374 ProLeuGlnLeuAlaProGlyValLeuHisProAlaProProProAlaProProLeu 393

QY 331 GCGCGATGGCGCGCGCTTCTTCTCGGAGATTAGACGATCGGAGCGGAATT 272
 DB 394 ValGlnProSerProPro-----ValAlaArgAlaAlaProValCys-----GluThr 409
 QY 271 ATCTTTTTCACCTCTGTGCGAGTCTCTGAGTGGGGGCTTCCGGAAGAGGGGAG 212
 DB 410 ValProValHisProLeuProGlnGlyGluValGlnGlyLeuProProProProPro 429
 QY 211 CCGGAGTCTCAGAGCCCGCGCGCTGCGCTGCCCTCT 173
 DB 430 ProProLeu-----ProProProGlyLeuArgProSer 440

RESULT 9

US-10-001-873-50
 ; Sequence 50, Application US/10001873
 ; Publication No. US20020160388A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Recipon, Hervé
 ; APPLICANT: Chen, Sei-Yu
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Turner, Leah
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P.
 ; FILE REFERENCE: DEX-0275
 ; CURRENT APPLICATION NUMBER: US/10/001, 873
 ; PRIORITY FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 60/252,055
 ; PRIORITY FILING DATE: 2000-11-20
 ; PRIOR APPLICATION NUMBER: 60/252,496
 ; PRIORITY FILING DATE: 2000-11-22
 ; NUMBER OF SEQ ID NOS: 55
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 50
 ; LENGTH: 1134
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 US-10-001-873-50

Alignment Scores:

Pred. No.: 0.0145 Length: 1134
 Score: 144.00 Matches: 53
 Percent Similarity: 37.97% Conservative: 7
 Best Local Similarity: 33.54% Mismatches: 64
 Query Match: 2.62% Indels: 34
 DB: 14 Gaps: 8

US-10-001-857-42 (1-3096) x US-10-001-873-50 (1-1134)

QY 559 CCATTATGCGCTACTTCTCCGAGTACACCGCGCGCGCGAGCACTGCGAGCGAG 500
 DB 976 ProLeuProLeuAlaArgLeuProProProProProProProPro----- 988
 QY 499 ATAAAGACCGCGCGCTCGCGCGCGCGCGCGCGCGCTCCCTCTCA-----GCCAGCGCGAC 446
 DB 989 ---LeuProAlaGProHisProProProProProProProProProProProPro 1007
 QY 445 CGGACGAGTGA-----CGCATGCGTATGCCGGGACCGGAGCGGAGCCGCGC 401
 DB 1008 GlnThrArgThrLeuProAlaAlaAlaArgThrMetAlaGlnPro-----ProPro 1023
 QY 400 CCCTTCCTTACTCCGGGAGCTTGCGGCTCCCGGCGACACCCCTTCTCTCGCGAGA 299
 DB 1024 ProArgLeuAlaLeuProArgArgArgArgProProArgProProProProProAla 1043
 QY 343 -----CGGCTTCGTTGCGCGATGCGCGCGCGCTTCTCTCGCGAGA 299
 DB 1044 ArgArgGlyProArgProThrProGlnAlaArgArgArgProArgProSerProArgArg 1063
 QY 298 TTAGAGAGATGCGAGACCGGAAGTATCTTTTTCACCTCTGTGCGAGTCTCGGATA 235
 DB 1064 Leu-----LeuArgSerPro-----HisSerLeuCysSerProArgLeu 1076

```

QY 238 GTGGGCGCTTCCCGAAGAG-----GCCGAGCCCGAGCTCTGAGCCGCCCGCTCT 185
Db 1077 ArgProGlyProAlaAspProArgArgGlnArgLaserThrsrProProArg 1096
QY 184 GCGCTGCCCTCTGGGAATTTCTTAGAAAGGTGACGGCAAGCGGAGAAACGG 131
Db 1097 SerIleProSerGlySerAlaCysArgProTPrArgThrGlyProArgSerPro 1114

RESULT 10
US-10-029-386-32198
/ Sequence 32198, Application US/10029386
/ Publication No. US20030194704A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
/ FILE REFERENCE: AEMICA-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ NUMBER OF SEQ. ID NOS: 34288
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 32198
/ LENGTH: 800
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC007663.28
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.3
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
US-10-029-386-32198

Alignment Scores:
Pred. No.: 0.0169 Length: 800
Score: 142.50 Matches: 76
Percent Similarity: 35.66% Conservative: 77
Best Local Similarity: 17.72% Mismatches: 74
Query Match: 2.59% Indels: 202
Db: 12 Gaps: 11

US-10-001-857-42 (1-3096) x US-10-029-386-32198 (1-800)
QY 1334 TCTGCAACGACGCTGCTCTCTTCTTAGTAAGGCTATAGACGCTGACGTAACACAGA 1275
Db 290 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 309
QY 1274 GTAAATTTCT-----ACTCTGCGAATACGCTAAACATGTTGGTGTCTAGTTCA 1224
Db 310 ILeThThThThThThThThThThThThThThThThThThThThThThThThThThTh 329
QY 1223 ACTCTGAGATCTCTCTCTCTCTCTGTCGACCTGACGACTCTTCTCTCTCTCTTGCA 1164
Db 330 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 349
QY 1163 TCATCTCCACATCTTTAGACATGCTGTAAGTGAAGATCTGTCACTGCTAGTACGATT 1104
Db 350 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 369
QY 1103 TTAATCCATTAAGTCAATGACGAAATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1044
Db 370 ILeThThTh-----ThThThThThThThThThThThThThThThThThThThThTh 384
QY 1043 TTTTCCCTTGCATGTCACAGATTTTCAAGATTTCCAGAGCAAAAGCCTTCATAGCAGA 984
Db 385 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 404
QY 983 TCTCTATAAGCTGTGATTATGATGATTAAGGACGTAATAACTCTCTGCTGCACTGTA 924
Db 983 TCTCTATAAGCTGTGATTATGATGATTAAGGACGTAATAACTCTCTGCTGCACTGTA 924

```

```

Db 405 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 419
QY 923 TGACCTTCAACACAGCTTATCAACAGCAAAACATGATCATATCCATTCAGTTCA 864
Db 420 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 439
QY 863 GCGAAGTGAGATCTTAAATTTAAATGATGCGATCTGTGATAGCTGTGCAAAATGAGA 804
Db 440 -----ThThThThThThThThThThThThThThThThThThThThThThThThTh 455
QY 803 ACTTTTCATTAAGTGTGTTCCATTCATGCGAGATCATTCATTCATTCATTTCA 744
Db 456 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 472
QY 743 ATAGACAGATGCTTCAAAAGACCAATAAGCTATCATGAGTAGTGTCCCAACTT 684
Db 473 ILeThThThThThThThThThThThThThThThThThThThThThThThThThThTh 492
QY 683 AATTCGCAAGCTCTTCAAAATCTGGGTAAGTGCACCGCACTGTGATTCCTTTTC 624
Db 493 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 512
QY 623 TCCATTTTTCGCAATCTGATCTCCATCCATCTGATGTCATCATACAGACGTTTC 564
Db 513 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 532
QY 563 ATACCATTAATGCTACTTCTCCCTGCTACCCGCGCGCCGACGCTGCGGACAC 504
Db 533 ILeThThThThThThThThThThThThThThThThThThThThThThThThThThTh 552
QY 503 G----- 503
Db 555 ThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 572
QY 502 -----GAATACGACGCGCT-----CG 483
Db 572 eThThThThThThThThThThThThThThThThThThThThThThThThThThThTh 592
QY 483 GCGCGCGCGCGCGCGCTCCCTCTGACCCGACCGGACCGGACGCTGACGATCGTAT 423
Db 592 CProProProProProProProProProProProProProProProProProProPro 612
QY 422 GCGCGGACGCGGACCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 378
Db 612 UProSerLeuProProProProProProProProProProProProProProProPro 632
QY 377 TTCGCGCTCCCGGACACCCCTCCCTCG-----CCACCGCTCGCTTCCGCGCATGGCG 321
Db 632 CserLeuProProProProProProProProProProProProProProProPro 652
QY 320 GCGCGCTTCTCTCTGCGGATTAAGACGATGCGGACGCGGAGTATCTTTTCA 261
Db 652 CProProSerLeuProProSerProProSerLeuSerProProSerLeu----- 668
QY 260 CCTCTGCGAGTCTGGTAGTGCGGCTTCCGGAAGAGCGGACCGCGAGTCTC 201
Db 669 -ProLeu-----ProProProProProProSerLeuPr 678
QY 200 AGAGCGCGCGCGCTGCGTGGCC 176
Db 678 CProProProProSerLeuProPro 686

RESULT 11
US-10-086-464-5
/ Sequence 5, Application US/10086464
/ Publication No. US20020199218A1
/ GENERAL INFORMATION:
/ APPLICANT: GORING, Daphne R. et al.
/ TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES
/ FILE REFERENCE: P 25,762-A USA
/ CURRENT APPLICATION NUMBER: US/10/086,464
/ PRIOR APPLICATION NUMBER: US 10/069,304

```

PRIOR FILING DATE: 2002-02-19
 PRIOR APPLICATION NUMBER: PCT/CA00/00966
 PRIOR FILING DATE: 2000-08-18
 PRIOR APPLICATION NUMBER: US 60/149,466
 PRIOR FILING DATE: 1999-08-19
 PRIOR APPLICATION NUMBER: US 60/159,122
 PRIOR FILING DATE: 1999-10-13
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 5
 LENGTH: 721
 TYPE: PRT
 ORGANISM: Brassica napus
 US-10-066-464-5

Alignment Scores:
 Pred. No.: 0.0178 Length: 721
 Score: 142.00 Matches: 58
 Percent Similarity: 37.04% Conservative: 22
 Best Local Similarity: 26.85% Mismatches: 75
 Query Match: 2.58% Indels: 61
 Gaps: 10

US-10-001-857-42 (1-3096) x US-10-066-464-5 (1-721)

QY 788 TGGTTTCATCATGCGCATCCATCTTGGATCCATCATTTCAATAGACAGCATGGCT 729
 Db 4 TtpSerProCysLeuSerSerProAlaLaserPheHisSerLeu----- 19
 QY 728 TCAAAAAGACCAATAGCTTATCATGATAGTTCCTCCCACTTATTTCTGCAGACGT 669
 Db 20 -----HisLeuProPhePheArgSerLeuAla 29
 QY 668 TCTTCAAAATCTTGGGTATGTCACCCAGCTTGTATGCTTTTCCATTTTCTGCG 609
 Db 30 AspMetSerSer-----AlaProSerProGlyThrGly 40
 QY 608 ATACTGAGCTCCCATCTGTAATGTCATCATCTACA-----GAACCTTTCATACCAT 555
 Db 41 SerProProSerProProSerAmSerThrThrThrProProProAlaSerAlaPro 60
 QY 554 ATGCTTACTTCTCCGCTGTCACCGCGCGCGCACGACATGTCGGAC----- 506
 Db 61 ProProThrThrProSer-SerProProProProSerThrLeuProThrSerProProPr 80
 QY 505 -ACGGAATACGACGCGCGCTGCGCGCGCGCGCGCGCGCGCTCCCTCCAGCCACGCGCA 447
 Db 80 oSerSerArgSerThrProSerAlaProProProProProProThrProSerThrProG 100
 QY 446 CCGGACAGCTGACGATGCGGTATGCCGAGACGCGGACCGCGC-----CCCTCTCCCT 390
 Db 100 ySerPro-----ProProLeuProGlnProSerProProAlaProTh 114
 QY 389 ACTCCGCGGAGACTTGGCGCTCCCGCGCACACCCCTCCCTCCAGCGCTCCGTC-- 332
 Db 114 rThProGlySerProAlaProValThrProProThrArgAsnProProProSerVal 134
 QY 331 -GGCCGATGCGCGCGCTTCTCTTCGCGAGATTAGAGACGATCGGAGACCGGA 276
 Db 134 lProGly-----ProProSer----- 139
 QY 275 AGTTATCTTTTTCACCTCTGTGCGAGTCTCTGGTATGGCGGCTTCCGGAAGAGGC 216
 Db 140 -----AsnProSerArgGlyGly-----GlySerProArgProProse 152
 QY 215 GGAAGCCGAGATCAGAGCCCGCGCGCTGCGCTCCCTCGGG 170
 Db 152 rSerProSerProProSerProSerSerArgGlyLeuSerThrGly 167

RESULT 12
 US-10-142-515-11
 Sequence 11, Application US/10142515
 Publication No. US20030078399A1

GENERAL INFORMATION:
 APPLICANT: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH
 APPLICANT: Lloyd, Kenneth O.
 APPLICANT: Yin, Beatrice W.T.
 TITLE OF INVENTION: Nucleic Acid Sequence Encoding Ovarian Antigen, CA125, and Use
 FILE REFERENCE: 649-A-US
 CURRENT APPLICATION NUMBER: US/10/142, 515
 CURRENT FILING DATE: 2002-07-23
 PRIOR APPLICATION NUMBER: US 60/290,480
 PRIOR FILING DATE: 2001-05-11
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: Patent In version 3.1
 SEQ ID NO 11
 LENGTH: 5877
 TYPE: PRT
 ORGANISM: Human Being
 FEATURE:
 NAME/KEY: MISC FEATURE
 LOCATION: (1)..(5877)
 OTHER INFORMATION: Amino acid sequence of MUC16B
 US-10-142-515-11

Alignment Scores:
 Pred. No.: 0.0404 Length: 5877
 Score: 141.50 Matches: 148
 Percent Similarity: 31.86% Conservative: 76
 Best Local Similarity: 21.05% Mismatches: 213
 Query Match: 2.57% Indels: 213
 Gaps: 27

US-10-001-857-42 (1-3096) x US-10-142-515-11 (1-5877)

QY 2237 CTTCATTATCTCTCTCTGTCATTTGAGGCCATCGCGACGATCAATGTCATCA 2178
 Db 713 LeuThrLeuAspThrSerThrThrPheMetSerGlyThrHis-----Ser 727
 QY 2177 ACCATGCGTAAAGAGATTCAGAGATACCAATATATATAGTACTGTCATCTGT 2118
 Db 728 ThrAlaSerGlnGlyPheSerHisSerGlnMetThrAlaLeuMetSerArgThrProGly 747
 QY 2117 AGAGTTCCAAATTCAGAGCCACTTGAAGATCTGTATCATTAATGCGAAGTTATGT 2058
 Db 748 GluValProThrLeuSerHisProSer-----ValGluGluAlaSerSer 762
 QY 2057 GGACCCAGGTACCTTAAAGCCAAATGTCGCTTGGGTTCTCTTCAACAGCATGG 1998
 Db 763 AlaSerPheSerLeuSerSerProValMetThrSerSerSerProValSerSerThrLeu 782
 QY 1997 TGTGAAGCGCTGCATCAACTTCTGCTATTAATAATTAATGATCAATCTCATCTG 1938
 Db 783 ProAspSerThrLeuHisSerSerLeuProVal----- 793
 QY 1937 CAAGTGGCAATTCCTCAAGATATGACCAAGCTTATCTCTGAGCCCTGTATG 1878
 Db 794 -----ThrSerLeuLeuThr-----SerGlyLeuValValys 803
 QY 1877 TCCATGATCTGATATGATACAGATGAGGCAACAGAGTGTACAAAGAGCTGAT 1818
 Db 804 ThrThrGluLeuLeuGlyThrSerSerGlnProGluThrSerSerProProAsnLeuSer 823
 QY 1817 ACAGTCTTACGCTGATATATATAGTAGAGACTGGGGGAAACACCTCGAGAT 1758
 Db 824 SerThrSerAlaGluThrLeuAlaThrThrGluValThrThrAspThrGlyLeuGlu 843
 QY 1757 CTGACAAAGACCGAGATGATCTTTCACCATGCTTCATGATGATGATGATGATGAT 1698
 Db 844 MetThrAsnValValThrSerGlyThrThrHisGlySer-----ProSerSer 859
 QY 1697 TTTTGTATCC-----ACCAAGAAAGTGTGTGTATACAGAGATCTTGAAGAAGACAT 1644
 Db 860 ValLeuAlaSerSerValThrThrThrAlaThrSerSerMetGlyLeuThrTyProThr 879
 QY 1643 GTGACTGTCTCAATAATTC-----CAGAAAAATTCAGAGATCAATCATGT 1599

```

Db      880 GlyAspThrValLeuThrSerThrProAlaPheSerAspThrSerArgIleGlnThr 899
QY      1598 AATTTGTTAAATTCACACACACAGAGATTGTTTATTCTATCTAATTAATCTTGCAAAA 1539
Db      900 Lys----- 900
QY      1538 TAGTTCACCAATTCCTCCCTTTAATTAATTTTGCATATCGAGGAAAGGTAGTGAAGT 1479
Db      901 -----Ser 901
QY      1478 AGCCTCTGTTCAAGAGGTTCAAAACCATCATATTGATGATCTCTTTGTAGTA 1419
Db      902 LysLeuSerLeuThrProGly----- 908
QY      1418 TCATTCGGGCGCTGGATGCGATGCAATGAATTGAATGGCAGAAAGAGATTCGCT 1359
Db      909 -----LeuMetGluThrSerIleSerGluGluThrSerSer 920
QY      1358 GCTTGAACATCAATTTTGTAGCTTCTGCAACAGACAGCTGCTCTTTCTAGTAAGGCT 1299
Db      921 Ala-----ThrGluYsSerThrValLeuSerSerValProThrGlyAla 935
QY      1298 ATAGACCTCTCGTAACACAGAGTAATTTCACTCTGCTGAATACTGCTAAACATTCG 1239
Db      936 ThrThrGluValSerArgThrGlu----- 943
QY      1238 TGGGTCTCTAGTCACTTCTGATCTCTTCTTCTCTGCTGAGCTTGAGTACTCTTT 1179
Db      944 ---AlaIleSerSerSerArgThrSerIleProGlyProAlaGln----- 957
QY      1178 ACTCTCTTGTGATGTCATCTCCACATCT- -TTAGCATGCT 1137
Db      958 -----SerThrMetSerSerArgThrSerMetGluThrIleThrArgIleSerThrPro 975
QY      1136 GTAACCTGGAAGATCTGTCACTAGCTAGCATTTTAAATCCAAATCATCTAGCTGCAAAA 1077
Db      976 LeuThrArgIleGluSerThrAspMetAlaIle---ThrPro-----Lys 989
QY      1076 TCTTCTCTTCAAAAACAGACGCTTATTACTTTTCCCTTGCAATGTCACAGATTTTC 1017
Db      990 ThrGlyProSerGlyAlaThrSerGlnGlyThrPheThrLeuAspSerSerSerThrAla 1009
QY      1016 AAGATCCCGACAGCAAAAGCCTTCATAGCAGATCTCTATAAAGCTCGATTATGAAG 957
Db      1010 SerTrpPro-----GlyThrHisSerAlaThrThrGlnArgPhe 1022
QY      956 TAAAGCAGCTAAATATCTGTCTGTGCCAGT- -GAATGCCCTCT 915
Db      1023 ProArgSerValValThrThrProMetSerArgGlyProGluAspValSerThrProSer 1042
QY      914 AAC---CAGCTTTCAAAACGCAAAACATGTATCCATATCCCTATCAGTCAGCAAG 858
Db      1043 ProLeuSerValGluIlyAsnSerProProSerSerLeuValSerSerSerSerValThr 1062
QY      857 GTGAGATCTTAAATTTAATAGTCCATCTGATGCTGTTGCAAAATTAAGAACTTT 798
Db      1063 SerProSerProLeuYsSerThrProSerGlySerSerHisSerSer----- 1078
QY      797 CGATTAACTTGGTTTCCATCATGCGACATCCATTTGGATTCATCTTCAATTCAATAGA 738
Db      1079 -----ProValProValThrSerLeuPheThrSerIleMetMetLysAla 1093
QY      737 ---GACATG---GCTTCAAAAAGCAAAATAGCTTATCATGAAGTGTCTCCCAAC 687
Db      1094 ThrAspMetLeuAspIleSerLeuGluProGluThr-----ThrSerAlaProAsn 1110
QY      686 TTTAATTCGACAGAGCTTCTTCAAAAATTTGGGTAAATGCCACCCAGTTTATTCCT 627
Db      1111 MetAsnIleThrSerAspIleSerLeuAlaIleSerIleAlaThrThrGluThrGluAla 1130
QY      626 TTCTCATTTTTCGTCAGTACTGAGCTCCGATCCGATCGCATCCATCCTACAGAGCT 567

```

```

Db      1131 IleHisValPheGluAsnThrAlaIleSerHisValGluThrThrSerAlaThrGluGlu 1150
QY      566 TTCATTAACCATTAATGCT- - - - -ACTTCCCGCTGGTCAAC 531
Db      1151 LeuTrpSerSerSerProGlyPheSerGluProThrLysValIleSerProValValThr 1170
QY      530 CGCGCCGCCACGCACTGTCCGACCAAGAAATAAGACGCCGCTCGAGCCGCCGCC 471
Db      1171 -----SerSerSerIleArg 1175
QY      470 CCCTCCCTCTGACGCCAGCCGACCGGACGGTGCAGCATGCGTATGCCCGGAGACGG 411
Db      1176 AspAsnMetValSerThrThrMetProGly-----SerSerGlyIleThrArgIleGlu 1193
QY      410 CGACCCCGCCCTCTCCCTTAATCTCCGAGGACTTCGGCGTCCCGGCACACCCCTCC 351
Db      1194 IleGluSerMetSerSerLeuThrProGly----- 1203
QY      350 CTGGCCACCGCTCCGTTCCGCGCGGACGGCGCCGCTCTCTTCGCGAGATTAGAGAC 291
Db      1204 -----LeuArgGluThrArgThrSerGlnAspIleThrSer 1215
QY      290 GATCGGAGACCGGAGATTATCTTTTCAACCTCTGTGCGAGTCTGCGTAGTGGCGG 231
Db      1216 SerThrGluThrSerThrValLeuTrpLys-----MetProSerGlyAla 1230
QY      230 TTCCCGGAA 222
Db      1231 ThrProGlu 1233

RESULT 13
US-10-243-8
; Sequence 8, Application US/10243243A
; Publication No. US20030104442A1
; GENERAL INFORMATION:
; APPLICANT: Ylm, Beatrice O.
; APPLICANT: Ylm, Beatrice W.T.
; TITLE OF INVENTION: Nucleic Acid Sequence Encoding Ovarian Antigen, CA125, and Use
; FILE REFERENCE: 649-B
; CURRENT APPLICATION NUMBER: US/10/243,243A
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: PCT/US02/14768
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: US 60/290,480
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 5935
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(5935)
; OTHER INFORMATION: Amino acid sequence of WC16B
US-10-243-243A-8

Alignment Scores:
Pred. No.: 0.0406 Length: 5935
Score: 141.50 Matches: 148
Percent Similarity: 31.86% Conservative: 76
Best Local Similarity: 21.05% Mismatches: 266
Query Match: 2.57% Indels: 213
DB: 15 Gaps: 27

US-10-001-857-42 (1-3096) x US-10-243-243A-8 (1-5935)
QY      2227 CTCCTATATCTTCTCTCTGCGCATTTGAGAGCCATCGGACGACGATCAATGTTGACATCA 2178
Db      784 LeuThrLeuAspThrSerThrThrPheMetSerGlyThrHis-----Ser 798

```

QY	217	ACCAATGGTAAAGAAATTCACAGAGATACCAATATATGTAATATGATACATCGTCGATACGT	211:8
Db	799	ThralaSerGlnGlyPheSerHisSerClnMetThrAlaLeuMetSerArgThrProGly	818
QY	2117	AGAGTTCCAATTCAAAGCCACTTGAAAGTACTGATCATATATGCAAGGTTATGTATAA	2058
Db	819	GluValProThrLeuSerHisProSer-----ValGlnGluAlaSerSer	833
QY	2057	GGACCCAGATCACTAAACAGGCCAAATGTGCTGGGTGGTCTGTTTCAACAGCATGG	1958
Db	834	AlaSerPheSerLeuSerSerProValMetThrSerSerProValSerSerThrLeu	853
QY	1997	TGTGAGCGCTGCATCAACCTTCTCTCGCTATTAAATAAATGTCATAACTATCCTG	1938
Db	854	ProAspSerIleHisSerSerSerLeuProVal-----	864
QY	1937	CAAGGTGCAAATTCCTCAAGATATATGCAACCAAGCTATCTCTGTCGAGCCCTGTATG	1878
Db	865	-----ThrSerLeuLeuThr-----SerGlyLeuValLys	874
QY	1877	TCCATGATCTTCGATTAAGACTACAGAAATGGCCGAACACAGTACGATACAAAGAGTGCAT	1818
Db	875	ThrThrGluLeuLeuGlyThrSerSerGluProGluThrSerSerProAsnLeuSer	894
QY	1817	ACAGTCCCTTACCCCTGTGATATTATATATAGAGTACCTTGGGGAAACACTCGGAGGAT	1758
Db	895	SerThrSerAlaGluIleLeuAlaThrThrGluValThrThrAspThrGluIleuSerLeu	914
QY	1757	CTGACAAAAGACCGAAGTGCATCTTTCACCATGTTCTGCATGAGATGAGTCCAAAGCC	1698
Db	915	MetThrAsnValValThrSerGlyThrThrHisGluSer-----ProSerSer	930
QY	1697	TTTTTGTTATCC-----ACCAGGAAAGTGTTGTGTAACAGAGATCTTGAAAGACACAT	1644
Db	921	ValLeuAlaAspSerValThrThrLysAlaThrSerSerMetGlyIleThrTyxProThr	950
QY	1643	GGTGACTGTTCACTAAATTC-----CAAAAAATTCAGATACAAATGT	1598
Db	951	GlyAspThrAsnValLeuThrSerThrProAlaPheSerAspThrSerArgIleIleThr	970
QY	1598	AAATTTGTTAAATTCACAACCTCACAGACAGTTTATTCTATCTTAATCTTGCAAAA	1539
Db	971	Lys-----	971
QY	1538	TAGTTACACATTTCTCCCTTTATTATTTTGTGATACGAGGAAGTACGTGATGAT	1479
Db	972	-----Ser	972
QY	1478	AGCCTTGCTTCACAGGGGTTCAAAACCCATCATATGATGATCTCCTTTGTAGTA	1418
Db	973	LysLeuSerLeuThrProGly-----	979
QY	1418	TCATCTCGGCGCTGATGCATGATGACATGAAATGATATGATGAGCAAGAAAGATCTGCT	1355
Db	980	-----LeuMetGluThrSerIleSerGlnGluThrSerSer	991
QY	1358	GCTTGACCAATCAATTTTGAAGCTTCTGCACACAGCATGCTGCTCTTTCTTGAATAAGCT	1299
Db	992	Ala-----ThrGluLysSerThrValLeuSerSerValProThrGlyAla	1006
QY	1298	ATAAGCAGTCACTGATACACAGAGAAATTCACCTGCTGCTGAATACGTAAACATTGT	1238
Db	1007	ThrThrGluValSerArgThrGlu-----	1014
QY	1238	TGGTCTTCTAGTTCACTTGGATCTCTTTCTTCTCCTGTGCACCTTGCAGTACTCTTT	1179
Db	1015	---AlaIleSerSerSerArgThrSerIleProGlyProAlaGln-----	1028
QY	1178	ACTCTTCTTTGATGATCAATCCTCCACATCT-----TTTACATGCGCT	1133
Db	1029	-----SerThrMetSerSerAspThrSerMetGluThrIleThrArgIleSerThrPro	1046
QY	1136	GTAATCGAAGATCTGTACACATGTTAGGCATTTTAAATCCATAAGTCAATTACTGAAAA	1077

```

Db 1047 LeuThrArgGlyGlnSerThrAspMetAlaIle---ThrPro-----Lys 1060
QY 1076 TCCTCCCTCTCAAAACAGACGCTTATTATTACTTTTCCCTGGACATGCACGATTTTC 1017
Db 1061 ThrGlyProSerGlyAlaThrSerGlnGlyThrPheThrLeuAspSerSerThrAla 1080
QY 1016 AAGATTCCCGAGCAAAAGCCTTCATGACGAGATCTTCTTAAAGTCGATTAATGATG 957
Db 1081 SerTrpPro-----GlyThrHisSerAlaThrTrpGlnArgPhe 1093
QY 956 TAAAGGACGCTAAATACTGCTGTGGCCAGT-----GAATGGCCTTCT 915
Db 1094 ProArgSerValValThrTrpPrometSerArgGlyProGlnAspValSerTrpProser 1113
QY 914 AAC---CACGTTATCAACAACAGCAAAAACATGATTCATTAATCCCTATCAGTTCAGGCAAG 858
Db 1114 ProLeuSerValGlnIlyAsnSerProProSerSerLeuValSerSerSerValThr 1133
QY 857 GTGAGACTTAATTAATTAACTAGAGCCATCCCTGTGATACCTGTGTCAAAATGAGAACTTT 798
Db 1134 SerProSerProLeuTrpSerThrProSerGlySerSerHisSerSer----- 1149
QY 797 CGATTAACTTGTTCCATCATGCGACGACATGCATCTTGAGATCCATCATTTCAATAGCA 738
Db 1150 -----ProValProValThrSerLeuPheThrSerTrpIleMetMetLysAla 1164
QY 737 ---GACATG-----GCTTCAAAAAGACCAAAATGCTTATCATCAATAAGTAGTCCCAAC 687
Db 1165 ThrAspMetLeuAspAlaSerLeuGlnProGlnThr-----ThrSerAlaProAsn 1181
QY 686 TTTAATCTCGACACGCTTCTCAAAATCTTGAGTAAATGCCACCAAGTTTGATTTGCTT 627
Db 1182 MetAsnIleThrSerAspGlnSerLeuValAlaSerLysAlaThrTrpGlnIlyAla 1201
QY 626 TTCCTCCATTTTGTGGCATACTAGCTCCCATCCGATCCGATCATCATCAACGAAGCT 567
Db 1202 IleHisValPheGlnAsnThrAlaAlaSerHisValGlnThrTrpSerAlaThrGlnGlu 1221
QY 566 TTCATAAACCATTAATAGCT-----ACTTCTCCGTCGTACG 531
Db 1222 LeuTrpSerSerSerProGlyPheSerGlnProThrLysValIleSerProValValThr 1241
QY 530 CGCGCCCGGCACGCACTGTCCGGACACAGAAATAGACCGCGCTCGCGCGCGCGCGC 471
Db 1242 -----SerSerSerIleArg 1246
QY 470 CCTCCCTCTACAGCCCGACCGGACGCGACGTCAGCATGCGTATGCCGGGACGCG 411
Db 1247 AspAsnMetValSerThrThrMetProGly-----SerSerGlyIleThrArgIleGlu 1264
QY 410 CGACCCCGCCCTCTCCCTACTACCGGGGACTTGGCGTCCCGGCGACACGCCCTCC 351
Db 1265 IleGlnSerMetSerSerLeuThrProGly----- 1274
QY 350 CTCGCCACCGCTCCGTTGGCGCGATGCGGCGCCCTTCTTCTTCGCGAGATTAGAC 291
Db 1275 -----LeuArgGlnThrArgTrpSerGlnAspIleThrSer 1286
QY 290 GATCGCGAGACCGAGATTAATCTTTTTCACCTCTGTGCGAGACGTCGGAGTAGCGGG 231
Db 1287 SerThrArgGlnThrSerThrValLeuTrpLys-----MetProSerGlyAla 1301
QY 230 TTCCCCGGAA 222
Db 1302 ThrProGln 1304

RESULT 14
US-09-823-240-2
: Sequence 2, Application US/09823240
: Patent No. US20020048813A1
: GENERAL INFORMATION
: APPLICANT: Frank B. Gertler

```

RESULT 14
US-09-823-240-2
; Sequence 2, Application US/09823240
; Patent No. US200204813A1
; GENERAL INFORMATION:
; APPLICANT: Frank B. Gertler

/ APPLICANT: James E. Bear
 / APPLICANT: Jürgen Wehland
 / APPLICANT: Joseph Loureio
 / TITLE OF INVENTION: Methods and Products for Regulating Cell
 / FILE REFERENCE: M0656/7064 (HCL)
 / CURRENT APPLICATION NUMBER: US/09/823,240
 / PRIOR FILING DATE: 2001-03-30
 / PRIOR APPLICATION NUMBER: 60/194,564
 / NUMBER OF SEQ ID NOS: 11
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 2
 / LENGTH: 802
 / TYPE: PRT
 / ORGANISM: Mus musculus
 / US-09-823-240-2

Alignment Scores:
 Pred. No.: 0.0267 Length: 802
 Score: 140.00 Matches: 49
 Percent Similarity: 36.24% Conservative: 5
 Best Local Similarity: 32.89% Mismatches: 58
 Query Match: 2.55% Indels: 37
 Gaps: 6

US-10-001-857-42 (1-3096) x US-09-823-240-2 (1-802)

QY 535 TCACCCGCGCGCCGACGACTGTCGACCGAATAACGCGCCGCTCGCGCGG 476
 Db 532 SerProThrProGlnGlyLeuVal-----LeuGlyProProAlaProPro 546
 QY 475 CCGCCGCTCCCTCTCACGCCGACCGCGACGCGTGCAGCATGCGTATGCCCGG 416
 Db 547 ProProProProLeuProSerGlyProAlaTyrAlaSerAlaLeuProProProGly 566
 QY 415 AGCGCGACCGCGCCGCTCTCTACTCCGCGGAGCTTCGCGCTCCCGCCGACACCC 356
 Db 567 -----ProProProProProProLeuProProSerThrGly-----ProProPro 580
 QY 355 CCTCCCTCCGACCGCTCTCCGCGCGATGCGCGCGCCCTTCTCTCTCGCGAGATTA 296
 Db 581 ProProProProProProProLeuProAlaGlnAlaProProPro----- 595
 QY 295 GAGACGATCGCGACCGAAGTATCTTTTTCACCCCTCTGCGAGTCCGCGTAGTG 236
 Db 596 -----ProProProProProAlaProProLeuProAlaSerGlyIlePhe 610
 QY 235 GCGCGTTCGCGAAGAG-----GCGAGCGCCGGA 206
 Db 611 SerGlySerThrSerGlnAspAsnArgProLeuThrGlyLeuAlaAlaAlaIleAlaGly 630
 QY 205 GTCTCGACCGCGCCGCTGCGCGCTGCGCGCTGCGAATTCTTAGAAGGTGACCGCA 146
 Db 631 AlaAlaLeuAlaGlyValSerArgValGluAspGlySerPhe-----Pro 645
 QY 145 GCGCGAAGGAAACCGTGAAGGTTTG 121
 Db 646 GlyGly-GlyAsnThrGlySerVal 653

RESULT 15
 US-10-196-935A-4
 / Sequence 4, Application US/10196935A
 / Publication No. US20030082720A1
 / GENERAL INFORMATION:
 / APPLICANT: Lifton, Richard P
 / APPLICANT: Wilson, Frederick H
 / APPLICANT: Choate, Keith
 / APPLICANT: Ishikawa, Kazuhiko
 / APPLICANT: Nelson-Williams, Carole
 / TITLE OF INVENTION: COMPOSITIONS METHODS AND KITS RELATING TO TREATING AND DIAGNOSING
 / TITLE OF INVENTION: HYPERTENSION
 / FILE REFERENCE: 044574-5113

/ CURRENT APPLICATION NUMBER: US/10/196,935A
 / CURRENT FILING DATE: 2002-10-25
 / PRIOR APPLICATION NUMBER: US 60/306,084
 / PRIOR FILING DATE: 2001-07-17
 / NUMBER OF SEQ ID NOS: 6
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO 4
 / LENGTH: 1243
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-10-196-935A-4

Alignment Scores:
 Pred. No.: 0.0341 Length: 1243
 Score: 139.50 Matches: 61
 Percent Similarity: 35.24% Conservative: 13
 Best Local Similarity: 29.05% Mismatches: 63
 Query Match: 2.54% Indels: 73
 Gaps: 10

US-10-001-857-42 (1-3096) x US-10-196-935A-4 (1-1243)

QY 535 TCACCCGCGCGCCGACGACTGTC-----CGACCGAATAACGCGCCGCTCGG 482
 Db 4 SerProAlaThrGlnThrThrValLeuMetSerGlnThrGlnAlaAspLeuAlaLeuArg 23
 QY 481 CCGCGCGCGCCG-----CCTCCCTCTCACGCCGACCGCCGCGCGCGAG 440
 Db 24 ProProProProLeuGlnGlyThrAlaGlyGlnProArgLeuGlyProProProArgAla 43
 QY 439 CGT-----GACGCGATGCGTATGCCCGGAGCGCC----- 410
 Db 44 ArgArgPheSerGlyLysAlaGluProArgProArgSerAlaGluSerArgArgSer 63
 QY 409 -----GACCGCGCCGCTCTCCGCGCGCTCTCCGCGCGCGCTCTCC 392
 Db 64 SerValAspLeuGlyLeuLeuSerSerThrSerLeuProAlaSerProAlaProAspPro 83
 QY 391 -----TTACTCCGCGGACTTCGCGCTCTCCGCGCGCGCTCTCC 350
 Db 84 ProAspProProAspSerAlaGlyProGlyProAlaArgSerPro-----ProProSer 101
 QY 349 TCGCACCGCGCTCGGTCGCGCGATGCGCGCGCTCTCTCTCCGCGAGATTGACAGC 290
 Db 102 SerGlyGluProProGlnGlyLysThrThrThrGlnGlyAlaProValLysAlaAlaGluAsp 121
 QY 289 ATCGCGAGACCGGAAGTATCTTTTTCACCCCTCTGCGAGTCCGCGTAGTGCGCGT 230
 Db 122 SerAlaArgProGlu-----LeuProAspSerAlaValGlyProGly 135
 QY 229 TCCCGAAGAGGCGGAGCCCGAGTCTCAGAGCCCGCGCTGCGCGCTCTCGG 170
 Db 136 SerArgGlu-----ProLeuArgValProGluAla 145
 QY 169 AATTCTTAGAAGGTAGCGGAGCGGAAGGAAACCGTGAAGTTGCGTAGACCTTA 110
 Db 146 ValAlaLeuGlnAlaGArgArgGlu----- 153
 QY 109 GTGCTGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 50
 Db 154 -----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 169
 QY 49 GACACGAGATTATG---TTTCTATACAG 23
 Db 170 AspGlyArgGlyLeuLysPheAspIleGlu 179

Search completed: November 25, 2003, 03:26:56
 Job time : 155 secs

Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

QY	1381	CATATATTCATGCATCATGCGCATCGAGGCCAGAAATGATATACATCAAAAGAGATATATCA	1440
Db	849	CACACCTCATTCACACACGCGCATCCAGGCTCAGAAATGSCATACCAAAAGAGACATATCA	908
QY	1441	ATTATGATGGGTTTGGAAACCCCTTGTAACGAGAGGCTACTCCACTACCTCCCTCGA	1500
Db	909	ATTATGATGGGTTTGGAGCCCTTGTAACGAGATACCTCCACCCACGTTCCCTCGC	968
QY	1501	TATGC	1505
Db	969	TATGC	973

RESULT	3
BM979022/c	
LOCUS	BM979022
DEFINITION	722 bp mRNA linear EST 21-FEB-2001
	UI-CF-DU1 adl-c-13-0-UI s1 UI-CF-DU1 Homo sapiens CDNA clone
	UI-CF-DU1 adl-c-13-0-UI s1, mRNA sequence.

ACCESSION	BM579022	GI:19599047
VERSION	BM579022.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 722)
Bonald, M.F., Lennon, G., and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCracken, B.

University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@iowa.edu
Tissue Procurement: Dr. M. V. Welch, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA sequencing by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=yes.

```

FEATURES
Source
Location/Qualifiers
1..722
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-adl-c-13-0-Ur"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1b="UI-CF-DUI"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT73-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG_lib=UI-CF-DUI

```

BASE COUNT	215	a	147	c	124	g	236	t
ORIGIN	TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368							
	TAG_SEQ=GCCTGTAGGC"							
Query Match	92.8%	Score	705	DB	12	Length	722	
Best Local Similarity	29.8%	Pred. No.	7,2e-117					
Matches	716	Conservative	0	Mismatches	0	Indels	1	Gaps
								1

QY	2380	CAAAAGTCGTAAACCGAAGCTTTAGCTTGATAGTGAACAGTGGCGGTATGAAACACAGGTT	2437
Db	722	CAAAAGTCGTAAACCGAAGCTTTAGCTTGATAGTGAACAGTGGCGGTATGAAACACAGG-T	664
QY	2440	TGCTCCATTCAACGAGTGTAGTGAACCCCGCGCGCAGTGACACTACTTACAGTTCGAAGGAAT	2495
Db	663	TGCTCCATTCAACGAGTGTAGTGAACCCCGCGCGCAGTGACACTACTTACAGTTCGAAGGAAT	604
QY	2500	GTCGACCTCATTAATAATATAGCCCTCTCTCTCACTCTCTCTGAAGCTGATGTGGCAGCTAG	2555
Db	603	GTCGACCTCATTAATAATATAGCCCTCTCTCTCTCACTCTCTGAAGCTGATGTGGCAGCTAG	544
QY	2560	TAAACACTTTTCAACAGCGCAAAAATATATGGAAAAATATTCCTTAACCCGCGACATGAGGT	2615
Db	543	TAAACACTTTTCAACAGCGCAAAAATATATGGAAAAATATTCCTTAACCCGCGACATGAGGT	484
QY	2620	TAAATGAAATTTTAAAGGTGGCCAAACCACTTTGTGGTTATGAAAGTATTATGGCAGAGAG	2675
Db	483	TAAATGAAATTTTAAAGGTGGCCAAACCACTTTGTGGTTATGAAAGTATTATGGCAGAGAG	424
QY	2680	ACACAAAAGGAATCTAAAGTTCTCTCTGAAATTTGATTTCTCTGCTCATTAATATTTTCC	2735
Db	423	ACACAAAAGGAATCTAAAGTTCTCTCTGAAATTTGATTTCTCTGCTCATTAATATTTTCC	364
QY	2740	TGTTGTGAAATCTGTTTGAAGACACTGCGGAGGCGGACATTAAGGCGCAGAGCTTCTCT	2795
Db	363	TGTTGTGAAATCTGTTTGAAGACACTGCGGAGGCGGACATTAAGGCGCAGAGCTTCTCT	304
QY	2800	TCAGACCCCACTCTTAGAGGGCAGACATCACAGGCTCCACATCACCGGAAGTGAGATGAT	2855
Db	303	TCAGACCCCACTCTTAGAGGGCAGACATCACAGGCTCCACATCACCGGAAGTGAGATGAT	244
QY	2860	TTCTTGCGGTACAACACTCATTTAAGGAATACCTTTAGTTTGAACAGCCCTATATGACATGA	2915
Db	243	TTCTTGCGGTACAACACTCATTTAAGGAATACCTTTAGTTTGAACAGCCCTATATGACATGA	184
QY	2920	ATGAAAACCTGCTGTTTAAAGTGGTTTATATATGTTCCATGGAAGAAGACTGCTTATTTGA	2975
Db	183	ATGAAAACCTGCTGTTTAAAGTGGTTTATATATGTTCCATGGAAGAAGACTGCTTATTTGA	124
QY	2980	ATGCATTTATATAGCTTATATGCTTTTATACAGATTTAATACAGAAATCATTTTTATATGA	3035
Db	123	ATGCATTTATATAGCTTATATGCTTTTATACAGATTTAATACAGAAATCATTTTTATATGA	64
QY	3040	ATGATTGATGAAATAATAGTGTATTAAAGGTTAATAAATTTCTTGACACAAAAAATAA	3096
Db	63	ATGATTGATGAAATAATAGTGTATTAAAGGTTAATAAATTTCTTGACACAAAAAATAA	7
RESULT 4			
LOCUS	BF969365	854 bp	mRNA linear EST 22-JAN-2001
DEFINITION	602271630P1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4359702 5',		
ACCESSION	BF969365		
VERSION	BF969365.1	GI:12336580	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 854)		
	NIH-MGC http://mgc.nci.nih.gov/ .		
	National Institutes of Health, Mammalian Gene Collection (MGC)		

JOURNAL COMMENT

Unpublished
Contact: Robert Strausberg, Ph.D.
Email: c9abds-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://lml.nhl.gov
Plate: LLM9999 row: 1 column: 07
High quality sequence stop: 785.
Location/Qualifiers

FEATURES

source

1. 854
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4359702"
/issue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1b="NIH MGC 84"
/note="Organ: adrenal gland; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."
BASE COUNT 273 a 158 c 186 g 237 t
ORIGIN

Query Match 22.5%; Score 696.8; DB 10; Length 854;
Best Local Similarity 94.5%; Pred. No. 2.1e-115;
Matches 809; Conservative 0; Mismatches 37; Indels 10; Gaps 8;

2245 GAAGGCGCTAGTATGAAAAAACAAG-AAAAAAGAAAGTGGCCCATGACCCGAG 2303
1 GAAAGGCGCTAGTATGAAAAAACAAGAAAGAAAGTGGCCCATGACCCGAG 60
2304 AATACATGAGCCAGCATATCAGAAATGTGCTGGATGTTTAAACATGTGAG 2363
61 AATACATGAGCCAGCATATCAGAAATGTGCTGGATGTTTAAACATGTGAG 120
2364 CATTGACATGAGCGCAAGTACGTAACCGAAGTTGAGCTGATGTAACAAGTTC 2423
121 CATTGACATGAGCGCAAGTACGTAACCGAAGTTGAGCTGATGTAACAAGTTC 180
2424 GGTATGACACAGGTTTGTCTCATTCAACAGTGTATGACCCGCGCCAGTCACT 2483
181 GGTATGACACAGGTTTGTCTCATTCAACAGTGTATGACCCGCGCCAGTCACT 240
2484 TACAGTTCAAGAAATGTCTGACCTCAATTAATATAGCCCTCTCTGAGTCTCTGAC 2543
241 TACAGTTCAAGAAATGTCTGACCTCAATTAATATAGCCCTCTCTGAGTCTCTGAC 300
2544 TGTATGTGCGACCTAGTAAGCACTTTCAACAGCAAAAATGATATGAAAAATATTCCTA 2603
301 TGTATGTGCGACCTAGTAAGCACTTTCAACAGCAAAAATGATATGAAAAATATTCCTA 360
2604 ACCCGGACCATGAGGTATATAGATTTTAAAGGTGCCAACCACCTTTGGGTATGA 2663
361 ACCCGGACCATGAGGTATATAGATTTTAAAGGTGCCAACCACCTTTGGGTATGA 420
2664 AGTATTTGGCAGGAGACACAAAAAGAAATCTAAAGTTCTCTGATTTGATTTCTGTG 2723
421 AGTATTTGGCAGGAGACACAAAAAGAAATCTAAAGTTCTCTGATTTGATTTCTGTG 480
2724 CTCAATTAATATTTCTGTGTAAGTGTGTTGAGAGACTGGGAGGTGGCCATTA 2783
481 CTCAATTAATATTTCTGTGTAAGTGTGTTGAGAGACTGGGAGGTGGCCATTA 540
2784 GGGGAGAGTCTTTCTTTCAGACCACTCTTAGAGGACATCAACAGGCTCCACATGAC 2843
541 GGGGAGAGTCTTTCTTTCAGACCACTCTTAGAGGACATCAACAGGCTCCACATGAC 600

2844 GGGAGTGCATGATTTTCTGGTAACTATATAGGAATACTTTAG-TTTGAC 2902
601 GGGAGTGCATGATTTTCTGGTAACTATATAGGAATACTTTAGTTTGAC 660
2903 AGCCTTATATACA-TGAATGAAACGCGTTTAA-AGTGGTTTATATGTCAT-G 2959
661 AGCCTTATATACA-TGAATGAAACGCGTTTAA-AGTGGTTTATATGTCATG 720
2960 GAAGAACTGCTCTTATGAAATGCAATGATGAAAGTATATGTTTATACAGATTAA 3019
721 GACAAACTGCTCTTATGAAATGCAATGATGAAAGTATATGTTTATACAGATTAA 778
3020 TCACAAATCATTTTATGAAATGCAATGATGAAAGTATATGTTTATACAGATTAA 3079
779 -TCCCAATCATTTTATGAAATGATG-TGTCACATGCTGTTTAACTGATTAATCTTG 835
3080 TCTTGACAAAAAAA 3095
836 CGCAAAAAACAAAAA 851
RESULT 5
CB169246/c
LOCUS CB169246 814 bp mRNA linear EST 30-JAN-2003
DEFINITION VBB603020215.R1 CSEQFXN41 testes Bos taurus CDNA, mRNA sequence.
VERSION CB169246.1 GI:28155372
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 814)
Adelson, D. L. and Gill, C. A.
Bovine ESTs (Adelson and Gill)
JOURNAL
TITLE Unpublished
AUTHORS Contact: David L. Adelson
COMMENT Animal Breeding and Genetics
Texas A&M University
Animal Science Dept., TAMU-2471, College Station, TX 77843-2471, USA
Tel: 9798452616
Fax: 9798456970
Email: david.adelson@tamu.edu.
FEATURES
source
1. 814
Location/Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/issue_type="testes"
/clone_1b="CSEQFXN41 testes"
/note="Organ: testes; Vector: pBluescript SK+; Site_1: NotI; Site_2: EcoRI; sequence 5' of the insert (5'-NNN-NNNinsert)
GCAATTTGAGCTCCACCGCGCGCGCGCGCGCGCGCTCGAG. Sequence 3' of the insert (AAGATTTGATATCAAGCTTATGATACCGTCACTCGAG, normalized Rd 1 library, sequenced 3' with M13R primer."
BASE COUNT 254 a 161 c 173 g 226 t
ORIGIN
Query Match 21.3%; Score 658.6; DB 14; Length 814;
Best Local Similarity 90.3%; Pred. No. 1.e-108;
Matches 758; Conservative 0; Mismatches 54; Indels 27; Gaps 4;
1348 ATGGTTCAACAGACATCTTTCTTGTGCAATCATATTCATGATCATGCAATCCAG 1407
814 ATGGTTCAACAGACATCTTTCTTGTGCAATCATATTCATGATCATGATCCAG 755
1408 GCCCAGAAATATCT-ACAAAAGAGATATCTCAATTTATGATGGTTTGAACCCCTGTG 1466
754 GCCCAGAAATATCT-ACAAAAGAGATATCTCAATTTATGATGGTTTGAACCCCTGTG 695

QY 1467 GAACGAGAGGCTACTTCCACCTACCTCCCTGATATGCAAAATATTAATTAAGGGAAGA 1526
 Db 694 TAAACCAAGAGCTACTTCCACCTACCTCCCTGATATGCAAAATATTAATTAAGGGAAGA 635
 QY 1527 AATGTGTAATTTTGGCAAGATTATAGATAGATATTAATTAAGGGAAGA 1586
 Db 634 AATGTGTAATTTTGGCAAGATTATAGATAGATATTAATTAAGGGAAGA 575
 QY 1587 TTTAAGCAATTTTCAATTTGATCCGATTTTCTGTGAAATTTAGTGAACAGTCCAGT 1646
 Db 574 TTTAAGCAATTTTCAATTTGATCCGATTTTCTGTGAAATTTAGTGAACAGTCCAGT 515
 QY 1647 TGTTCCTTCAAGATCTCTGTATACCAACCACTTCTGTGATATGCAAAATTTTGG 1706
 Db 514 TGTTCCTTCAAGATCTCTGTATACCAACCACTTCTGTGATATGCAAAATTTTGG 455
 QY 1707 AATCATCTATGCAAGATGGAAGATGCACTTCTGTGATATGCAAAATTTTGG 1766
 Db 454 AATCATCTATGCAAGATGGAAGATGCACTTCTGTGATATGCAAAATTTTGG 396
 QY 1767 GTGCTTCCCCCAAGTCTACTATATATATATATATATATATATATATATATATATAT 1826
 Db 395 -TGCTTCCCCCAAGTCTACTATATATATATATATATATATATATATATATATATAT 337
 QY 1827 TTTGTCTACTACTGT 1886
 Db 336 TTTGTCTACTACTGT 277
 QY 1887 GCTGCAAGAGAGATTAAGT 1946
 Db 276 GCTGCAAGAGAGATTAAGT 219
 QY 1947 TTTATGCAATTTTATTTTATAGGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2006
 Db 218 -----AGGCAAGAGAGT 180
 QY 2007 TGAACAG 2066
 Db 179 TGAACAG 120
 QY 2067 ACCTTGCAATTTATATACAGTCTTTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2125
 Db 119 ACCTTGCAATTTATATACAGTCTTTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
 QY 2126 CACGAGTACTATTTATATATTTGTATCTCTGATATTTCTTTACGATGATGTGTC 2184
 Db 59 CACGAGTACTATTTATATATTTGTATCTCTGATATTTCTTTACGATGATGTGTC 1
 RESULT 6
 AK050805
 LOCUS 3071 bp mRNA linear HTC 05-DEC-2002
 DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
 enriched library, clone: D030020M24 product: CORNELL MOUND HEALING
 RELATED PROTEIN homolog (Rattus norvegicus), full insert sequence.
 ACCESSION AK050805
 VERSION AK050805.1 GI:26094130
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 COMMENT
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, H., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, Y., Nishi, K., Kitanaka, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kite, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multiplexed sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL 20530913
 MEDLINE 11076861
 PUBMED
 REFERENCE
 AUTHORS Kawai, J., Shibata, K., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, D., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J.,
 Schriml, L., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T.,
 Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barish, G.,
 Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F.,
 Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M.,
 Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M.,
 Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J.,
 Mombere, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S.,
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 JOURNAL 21085660
 MEDLINE 11217851
 PUBMED
 REFERENCE
 AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 JOURNAL 6 (bases 1 to 3071)
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koyu, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnishi, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku-Akashira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toy, T., Yasunishi, A.,
 Muramatsu, M., and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-research.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT
 The cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.


```

Db      181  GAACTGTGCTGAGATGTTTAAACCATGATGAGCTTTGACATGACGCGCAAGTACG 240
Qy      2389 TAAACCGAAGTTGAGCTTATGATGACAAAGTTGGTATGAAACACAGGTTGCTCCATT 2448
Db      241  TAAACCGAAGTTGAGCTTATGATGACAAAGTTGGTATGAAACACAGGTTGCTCCATT 300
Qy      2449 CAACAGTGTATGACCCCGCGCAGTGCACACTACTACAGTCAAGAAATGCTGACCT 2508
Db      301  CAACAGTGTATGACCCCGCGCAGTGCACACTACTACAGTCAAGAAATGCTGACCT 360
Qy      2509 CAATAA-ATAATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2567
Db      361  CAATTAACATATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Qy      2568 TTCAACAGGCAAAATATATTTGAAATATCTCTTAACCCGACCATGAGTTATGAA 2627
Db      421  TTCAACAGGCAAAATATATTTGAAATATCTCTTAACCCGACCATGAGTTATGAA 480
Qy      2628 TTTTAAAGTTGCCAAACCAACTTGTGTATGAAATTATTTGCGACGACGACGACAAA 2687
Db      481  TTTTAAAGTTGCCAAACCAACTTGTGTATGAAATTATTTGCGACGACGACGACAAA 540
Qy      2688 AGGATCTAAAGTCCCTGGAATTTGATTTCTCTGCTCATTAATATTTTCTGTTGTA 2747
Db      541  AGGATCTAAAGTCCCTGGAATTTGATTTCTCTGCTCATTAATATTTTCTGTTGTA 600
Qy      2748 AACTTTGTTAGAGAGACTGGGAGGTGGCCAT-AAAGGGGACAGTCTTCTTCAGACC 2806
Db      601  AACTTTGTTAGAGAGACTGGGAGGTGGCCAT-AAAGGGGACAGTCTTCTTCAGACC 660
Qy      2807 CAATCTTAAGGGCAATGACCGAGGCTCCACATGACGGAAGTGAAGTGAATTTCTGG 2866
Db      661  CAATCTTAAGGGCAATGACCGAGGCTCCACATGACGGAAGTGAAGTGAATTTCTGG 719

RESULT 8
BO965085      957 bp      mRNA      linear      EST 21-AUG-2002
LOCUS      AGSCOURT 10052197 NIH_MGC_134 Mus musculus cDNA clone
DEFINITION      IMAGE:6509503 5', mRNA sequence.
ACCESSION      BO965085
VERSION      BO965085.1 GI:22380563
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM14076 row: 9 column: 08
High quality sequence stop: 632.
Location/Qualifiers
1..957
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6509503"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_134"
/notes="Vector: pCMV-Sport6.1.cdb; Site_1: EcoRV; Site_2:

```

```

BASE COUNT      274 a      198 c      253 g      232 t
ORIGIN
Query Match      20.9%; Score 648; DB 13; Length 957;
Best Local Similarity 87.7%; Pred. No. 1.3e-106;
Matches 719; Conservative 0; Mismatches 100; Indels 1; Gaps 1;
NotI: Cloned unidirectionally. Primer: Oligo dT. Average
insert size 1.7 kb. Constructed by ResGen, Invitrogen
Corp. Note: this is a NIH_MGC Library."

Qy      541  GGAGAAGTGGATTAATGTTATGAAAGCTTCTGATGATGATGATGATGATGATGATGATG 600
Db      26  GGCGGGAGCGGATATGTTATGAAAGCTTCTGATGATGATGATGATGATGATGATGATG 85
Qy      601  CTCAGTATGCGCAAAATATGAGAAAGCAATACAACTGGGTGACATTACCCAGAT 660
Db      86  CTCGGGGTCCCGGAAAAATGAGAAAAAGTACACAACTGGGTGACATTACCCAGAT 145
Qy      661  TTTGAAGAGCTTGTGCGAATTTAAAGTTGGAGAACTCTTCAATGATTAAGTATTGGT 720
Db      146  TTTGAAGAGCTTGTGCGAAGCTGAAAGTTGGAGAACTCTTCAATGATTAAGTATTGGT 205
Qy      721  CTTTGAAGCCATGCTGCTATTAATGATGATCCCAAGATGATGCTGCTGATGAT 780
Db      206  CTTTGAAGCCATGCTGCTATTAATGATGATCCCAAGATGATGCTGCTGATGAT 265
Qy      781  GGAAACCAAGTTAATGAAAGTTCTCAATTTTGAACAAGCTATGAGATGCGACTATT 840
Db      266  GGAAACCAAGTTAATGAAAGTTCTCAATTTTGAACAAGCTATGAGATGCGACTATT 325
Qy      841  AAAATTAAGATCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db      326  AAAATTAAGATCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 385
Qy      901  TTGATTAACGTGTGTAAGAGCCATTCAGTGGCAGACAGATTTTACGTCCTTACAT 960
Db      386  TTGATTAACGTGTGTAAGAGCCATTCAGTGGCAGACAGATTTTACGTCCTTACAT 445
Qy      961  CATATCCAGATTTTAAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db      446  CATATCCAGATTTTAAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 505
Qy      1021  ATCTGACATTCGCAAGGAAAAAGTAAATAAAGTGTCTGTTTGAAGAGCAAGATT 1080
Db      506  ATCTGACATTCGCAAGGAAAAAGTAAATAAAGTGTCTGTTTGAAGAGCAAGATT 565
Qy      1081  CAGTCAATGATTTATGATTTTAAATGCTTAACTGTGACAGATCTTGAAGTACAGGC 1140
Db      566  CAGTCAATGATTTATGATTTTAAATGCTTAACTGTGACAGATCTTGAAGTACAGGC 625
Qy      1141  ATGCTAAAGATGTGAGAGATGACATGCAAAAGAGTAAAGTAAAGTAAAGTAAAGTAA 1200
Db      626  ATGCTAAAGATGTGAGAGATGATGCAAAAGAGTAAAGTAAAGTAAAGTAAAGTAAAG 685
Qy      1201  GGAAGAAAGAGATGCAAGATTTGACTGAACACCAAGATTTTACAGATTTACAG 1260
Db      686  GGAAGAAAGAGATGCAAGATTTGACTGAACACCAAGATTTTACAGATTTACAG 745
Qy      1261  AGAGTAAATTTATCTGCTGTTTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1320
Db      746  AGAGTAAATTTATCTGCTGTTTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 804
Qy      1321  AGTGTGTTGCGAAGCTCAAAAATTGATGTTCAAGCAG 1360
Db      805  AGTGTGTTGCGAAGCTCAAAAATTGATGTTCAAGCAG 844

RESULT 9
A1790514/c      828 bp      mRNA      linear      EST 02-JUL-1999
LOCUS      A1790514
DEFINITION      U102b03.x1 Sugano mouse kidney mRna Mus musculus cDNA clone
IMAGE:2064845 3' similar to WP.T23B12.4 CE14032 //, mRNA sequence.
ACCESSION      A1790514

```


Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLY-A:Yes.

FEATURES

Location/Qualifiers
1..616
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-adr-h-18-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DUI"
/note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site_1: Ecor I; Site_2: Not I; UI-CF-DUI is a normalized cDNA library containing the following tissue(s): Primary Lung Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dri)18 tail. The sequence tag for this library is GGCTGAGGC.
TAG LIB=UI-CF-DUI
TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG SEQ=GGCTGAGGC"

BASE COUNT 190 a 123 c 98 g 205 t

ORIGIN

Query Match 19.7%; Score 610; DB 12; Length 616;
Best Local Similarity 100.0%; Pred. No. 9.6e-100;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2487 AGTTCAAGAAATGCTGACCTCATATATATAGCCCTCTCTGAGTCTCTGAACTGT 2546
DB 616 AGTTCAAGAAATGCTGACCTCATATATATAGCCCTCTCTGAGTCTCTGAACTGT 557
QY 2547 ATGGGAGAGTAGTAAGCACTTTCACAGGCAAAAATATATGGAATAATTCCTAAC 2606
DB 556 ATGGGAGAGTAGTAAGCACTTTCACAGGCAAAAATATATGGAATAATTCCTAAC 497
QY 2607 CGGACATGAGGTTAATGAATTTTAAAGGTTGCCAAACCACTTGTGTTATGAAGT 2666
DB 496 CGGACATGAGGTTAATGAATTTTAAAGGTTGCCAAACCACTTGTGTTATGAAGT 437
QY 2667 TATTGGAGAGGAGACAAAAGAAATCTAAAGTCTCTGAAATTTGATTTCTGCTC 2726
DB 436 TATTGGAGAGGAGACAAAAGAAATCTAAAGTCTCTGAAATTTGATTTCTGCTC 377
QY 2727 ATAAATATTTTCTGTTGTAACCTGTTTGAAGAGACTGGGAGAGTGCCATAAGGG 2786
DB 376 ATAAATATTTTCTGTTGTAACCTGTTTGAAGAGACTGGGAGAGTGCCATAAGGG 317
QY 2787 GCGAGAGCTCTTTCAGACCACTCTTAAGGGGACATACAGAGCTCCACATCAGGG 2846
DB 316 GCGAGAGCTCTTTCAGACCACTCTTAAGGGGACATACAGAGCTCCACATCAGGG 257
QY 2847 AAGTGAGATGATTTCTTGGGTAAACAATCATTAAGAAATCTTTTATGTTGACAGCC 2906
DB 256 AAGTGAGATGATTTCTTGGGTAAACAATCATTAAGAAATCTTTTATGTTGACAGCC 197
QY 2907 TTTATGAGATGAGTAAAGAACTGCTGTTTAAAGTGGTTTATATGTTTCAATGAAGAA 2966
DB 196 TTTATGAGATGAGTAAAGAACTGCTGTTTAAAGTGGTTTATATGTTTCAATGAAGAA 137
QY 2967 CTGGCTTTATGATGATGATGAACGTTATATGTTTATATGATGATTAATCAACA 3026
DB 136 CTGGCTTTATGATGATGATGAACGTTATATGTTTATATGATGATTAATCAACA 77

QY 3027 TCATTTTATGATGATGATGAAGAAATAGTGTATTAAGGTTAATAATTTCTGAC 3086
DB 76 TCATTTTATGATGATGATGAAGAAATAGTGTATTAAGGTTAATAATTTCTGAC 17
QY 3087 AAAAAAAAAA 3096
DB 16 AAAAAAAAAA 7

RESULT 11
CA977646
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CA977646 900 bp mRNA linear EST 06-JAN-2003
AGENCY: NIH-MGC, 164 Mus musculus cDNA clone
IMAGE:30146905 5', mRNA sequence.
CA977646.1 GI:27510300
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 900)
NIH-MGC <http://mgi.mc.man.ac.uk/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. David Rowe and Dr. Mina
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Place: NDAM0063 row: 6 column: 02
High quality sequence stop: 655.

FEATURES

Location/Qualifiers
1..900
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30146905"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 164"
/note="Vector: PCMV-SPORE6.1.cdb; Site 1: EcorV; Site 2:
NotI. Non-normalized full-length enriched library from
pooled mouse embryonic limb, maxilla and mandible, day
10.5 and 11.5 (site selected for the 0.5-1 kb fragments)
cloned directionally, priming method: Oligo-dT. cDNA
enrichment: >1k bp, Average insert size 1.8k bp. Priming
sequence: 5'-GACTAGTCTAGATCGCGCGCCGCCCT(T) 3'. Tissue
contributed by: David Rowe. Library constructed by Resgen,
Invitrogen Corp."

BASE COUNT 235 a 213 c 212 g 239 t 1 others

ORIGIN

Query Match 19.4%; Score 599.2; DB 14; Length 900;
Best Local Similarity 85.5%; Pred. No. 8.1e-98;
Matches 700; Conservative 0; Mismatches 94; Indels 25; Gaps 2;

QY 1535 ACTATTTTGCAGATTAATAGTAAATTAAGAACTGTGTGAGGTTGTAATTAACAA 1594
DB 3 ACTATTTTGCAGATTAATAGTAAATTAAGAACTGTGTGAGGTTGTAATTAACAA 62
QY 1595 ATTTCATTTGATCTGATTTTCTGTTGTAATTTAGTAAACAGTCACATGTCTTT 1654
DB 63 ACTTCACCTGATCTGATTTTCTGTTGTAATTTAGTAAACAGTCACATGTCTTT 122
QY 1655 CAAGATCTCTGTTCAAAACGATTTCTGCGGTAAACAAAGGCTTTTGAAGTCAATC 1714
DB 123 CAAGATCTCTGTTCAAAACGATTTCTGCGGTAAACAAAGGCTTTTGAAGTCAATC 182

DEFINITION	wt02f11.x1 NC1_CGAP_C03 Homo sapiens cDNA clone IMAGE:2506317 3'
ACCESSION	AF007593
VERSION	AF007593.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/hc/cgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
FEATURES	<p>CDNA Library Preparation: M. Bento Soares, Ph.D.</p> <p>CDNA Library Arraying: Greg Lennon, Ph.D.</p> <p>DNA Sequencing by: Washington University Genome Sequencing Center</p> <p>Clone distribution by: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLNM at: www.bio.ln1.gov/bbrip/image.html</p> <p>Insert Length: 715 Std Error: 0.00</p> <p>Seq primer: -40UP from Gibco</p> <p>High quality sequence stop: 459.</p> <p>Location/Qualifiers</p> <p>1..598</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:2506317"</p> <p>/sex="pooled"</p> <p>/tissue.type="colon"</p> <p>/lab.host="DH10B"</p> <p>/clone.lib="NCI CGAP C03"</p> <p>/note="Vector: pTR73-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization."</p>
BASE COUNT	169 a 122 c 98 g 188 t 1 others
ORIGIN	
Query Match	19.3%; Score 597; DB 9; Length 598;
Best Local Similarity	99.8%; Pred. No. 2.1e-97;
Matches 597; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	2490 TCAGGAATGCTCAGCTCATAATATATYAGCCCTCCTCAGCTCCTGACTGATG 254
Db	598 TCAAGGAATGCTCAGCTCATAATATATAGCCCTCCTCAGCTCCTGAACGTATG 539
OY	2550 TGGCAGCTAGTAGCATTTCACACGGCAAAATGATATGGAAATATTCCTAACCCGG 2608
Db	538 TGGCAGCTAGTAGAGCACTTCACACGGCAAAATGATATGGAAATATTCCTAACCCGG 479
OY	2610 ACCATGAGGTATATGAAATTTTAAAGGTGGCCAAACCACTTTGGTGTATGAAGTTAT 2665
Db	478 ACCATGAGGTATATGAAATTTTAAAGGTGGCCAAACCACTTTGGTGTATGAAGTTAT 419
OY	2670 TGGCAGAGGACACAAAAGGATCTAAAGTTCTCTGAAATTTGATTTCTGCTCAT 2728
Db	418 TGGCAGAGGACACAAAAGGATCTAAAGTTCTCTGAAATTTGATTTCTGCTCAT 359
OY	2730 AATATTTTCCGTGTGAACCTGTGTTAGAGAGACTGGGAGGTGGCCATTAAGGGCA 2788
Db	358 AATATTTTCCGTGTGAACCTGTGTTAGAGAGACTGGGAGGTGGCCATTAAGGGCA 299
OY	2790 GAGCTCTTTTCAGACCCCACTCTTAGAGGGACATCACAGCTCCACATCGGGAA 2849

Db		298	GAGCTCTTTCGACCCCAACTCTTAGAGGGCACATCACCAGGCTCCACATCCAGCGAAG	239
QY		2850	TGAGATGCATTTCTTGCGGTAAACAACCATATAAGAATCTTTAGTTTGACAGCCTTA	2909
Db		238	TGAGATGCATTTCTTGCGGTAAACAACCATATAAGAATCTTTAGTTTGACAGCCTTA	179
QY		2910	TATGACATGAATGAAAACTGCTCTTTTAAAGTGCTTTATTATGTTCCATGGAACAACTG	2969
Db		178	TATGACATGAATGAAAACTGCTCTTTTAAAGTGCTTTATTATGTTCCATGGAACAACTG	119
QY		2970	GTCCTATTGATGCATTTGATGAACGGTTATATGTTTTATTACAGATTTATTCACAAATCA	3029
Db		118	GTCCTATTGATGCATTTGATGAACGGTTATATGTTTTATTACAGATTTATTCACAAATCA	59
QY		3030	TTTTTATGATGATGATGATGAAAATAGTGTTTATATAAGGTTAATAAATTTCTTGACA	3087
Db		58	TTTTTATGATGATGATGATGAAAATAGTGTTTATATAAGGTTAATAAATTTCTTGACA	1
RESULT_14				
BUS058931				
LOCUS				
DEFINITION			718 bp mRNA linear EST 26-AUG-2002	
ACCESSION			U1-M-FR0-cam-e-22-0-U1.r1 NIH Mus_FRO Mus musculus cDNA clone	
VERSION			IMAGE:6414333 5', mRNA sequence.	
KEYWORDS			BUS058931	
SOURCE			BUS058931.1 GI:22499220	
ORGANISM			EST.	
			Mus musculus (house mouse)	
			Mus musculus	
			Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE			1 (bases 1 to 718) NIH-MGC http://mgc.nci.nih.gov/.	
AUTHORS			National Institutes of Health, Mammalian Gene Collection (MGC)	
TITLE			Unpublished	
JOURNAL			Contact: Robert Strausberg, Ph.D.	
COMMENT			Email: ggaaps-f@mail.nih.gov	
			Tissue Procurement: Dr. Jim Lin, University of Iowa	
			cDNA Library preparation: Dr. M. Bento Soares, University of Iowa	
			cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa	
			DNA Sequencing by: Dr. M. Bento Soares, University of Iowa	
			Clone Distribution: MGC clone distribution information can be	
			found through the I.M.A.G.E. Consortium/LNL at:	
			http://image.lnl.gov	
			This clone was contributed by the Brain Molecular Anatomy Project	
			(BMAP)	
FEATURES			Seq primer: pyX-5.	
source			Location/Qualifiers	
			1..718	
			/organism="Mus musculus"	
			/mol_type="mRNA"	
			/strain="C57BL/6"	
			/db_xref="taxon:10090"	
			/clone="IMAGE:6414333"	
			/tissue_type="whole brain"	
			/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"	
			/lab_host="DH10B (r1 phage resistant)"	
			/clone_lib="NIH BMAP FRO"	
			/note="Organ: Brain; Vector: pyX-Asc; Site 1: Ecov I;	
			Site 2: Not I; The library was constructed according	
			Bonaldo, Lennon and Soares, Genome Research, 6:791-806,	
			1996. Denatured RNA was size fractionated on a 1% agarose	
			gel. First strand cDNA synthesis was primed with a 16-mer	
			primer containing a Not I site. Double strand cDNA was	
			size selected according to rRNA size fraction. Ligated	
			with EcoR I adaptor, digested with NotI and then cloned	
			directionally into pyX-Asc vector. The library tag	
			sequence located between the Not I site and the poly A tail	
			is AGCGAGACAG. This library was created for the University	
			Iowa Brain Anatomy Project (BMAP). *Gene Discovery in the	
			Developing Mouse Nervous System", supported by National	
			Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,	

Tue Nov 25 09:35:52 2003

us-10-001-857-42.rst

Page 14

Db	601	AGGAATCTAAGTTCCTCCTGAA-TTGATTTCTCTGCTCATA---ATATTCTCTGTGTA	656
Qy	2748	AACTTGTGAGAGAGACTGGGAGG	2773
Db	657	AACTGTTGAGAGAGACTGGAGGTGG	682

Search completed: November 25, 2003, 02:50:11
Job time : 6201 secs

and is derived by analysis of the total score distribution.

XX

Human lung specific
Human molecule from
Drosophila melanogaster
Human secretory protein
Novel human secretory
Hedichidinically active
Novel human diatom
Hedichidinically active
Human transmembrane
MASP homolog protease
Amino acid sequence
Human tumour suppressor
Hedichidinically active
Human polypeptide
Human Scar1 protein
Amino acid sequence
Human pab10 (Pro-a)
Human WAVE1 (WAVE)
Human lung-specific
Streptomycetes virid.
Novel human diatom
Human novel polypep
Human diaphanous p
Streptomycetes virid.
S. cerevisiae BAX-9
Novel human diatom
Human mucin (MUC-1
Arabidopsis thaliana
Arabidopsis thaliana
Streptomycetes virid.
Mammalian Era (Men
Mouse neural Men+
Mouse neural Men+
Mouse neural Men+
Mammalian enabled
Human kinases and
Human MNK1 protein
Cotton fiber-specif
Np70 protein cabdo
Human Npw58BP1 tra
Np70 protein sequen
Human RNA metaboli
NpwBP. Homo sapi
Human colon cancer
Hypoxia-regulated


```

Db      480  iStrMetLeuLeuYsgInGluProGlnArgLnhIstLeuAlaCysLeuGlyThrTrpV 500
QY      2055  TCCTTACCATACCTTCGCTATATATATACAGTACTTCTTAAGTGGCTTGAATGGAAAC 2114
Db      500  aLeuYrHsEnLeuArgIleMeCileGlnTrLeuLeuSerGlyPheGluLeuGluL 520
QY      2115  TCTACAGTATGCAGAGTACTATATACATATATATGATCTCTGATACCTTACGCAT 2174
Db      520  euYrSerMetHsGlyTrpYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 540
QY      2175  GATTGATGTCACATTTAGTTCGCGCATGCTCTCAATTCGACAGAGAAAGCAATATG 2234
Db      540  rPheMetSerThrLeuSerArgAlaAspGlySerGlnMetIleGluIleGlyIleMeG 560
QY      2235  AAGACAGACAGAAAGCCGCTAGTAAATTAACAAAGAAAGAAAGAAAGTTCGCCCAT 2294
Db      560  lUGlInGlnGlnYsgIArgSerSerYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 580
QY      2295  TGAAGCCAGACATGACATGAGCCAGCATATCAGAACATGTGTCTGGAAATGTTTAAA 2354
Db      580  euSerArgGluIleThrMetSerGlnAlaTyGlnMetCysAlaGlyMetPheLysT 600
QY      2355  CCATGGTACATTTGACATGACGCGCAAGTACGTAAACGAGTTGAGCTTGAATGAG 2414
Db      600  hrMetValAlaPheMetLeuAspGlyYrValArgYrProYrPheGluLeuAspSerG 620
QY      2415  AACCAAGTTGGATATGACACAGGTTTCTCATTCACAGTGTGATGACCCCGCGCAG 2474
Db      620  lUGlInValArgYrGluHsArgPheAlaProPheMetSerValMetThrProProP 640
QY      2475  TGACACTACTACAGTTCAGAAATGCTGACCTCAATTAATATAGCCCTCTCTCACT 2534
Db      640  aHisTrpLeuGlnPheYrGlnMetSerAspLeuMetYrYrSerProProGlnS 660
QY      2535  CTCCTGAAGTATGATGACGAGTATGACACTTTCAACGAGCAAAATGATATGAGAA 2594
Db      660  erProGluLeuYrValAlaAlaSerYrHsPheGlnGlnAlaYrMetIleLeuGluA 680
QY      2595  ATATTCTTACCCGAGCCATGAGTAAATATTAAGTTGGCAAAACCAACTTTG 2654
Db      680  snIleProAsnProAspHsGlnValAsnArgIleLeuYrValAlaYrProAsnPheV 700
QY      2655  TGGTATGAGATTTATGACGAGAGACACAAAGCAATCTAAAGTCTCTGAAATTTG 2714
Db      700  aIValMetYrLeuLeuAlaGlyGlyHsIleYrYrYrYrYrYrYrYrYrYrYrYrYr 720
QY      2715  ATTTCTCTGCTCATTAATATTTTCTGTTGAAAATTGTT 2755
Db      720  spPheSerAlaHisYrYrPheProValValYrLeuVal 733

```

```

PR      09-FEB-2001; 2001US-268117P.
PR      15-FEB-2001; 2001US-269618P.
PR      23-FEB-2001; 2001US-271118P.
PR      07-MAR-2001; 2001US-274486P.
PR      09-MAR-2001; 2001US-274436P.
PR      28-NOV-2001; 2001US-314229P.
PR      01-FEB-2002; 2002US-353284P.
PR      (INCY-) INCYTE GENOMICS INC.
XX
XX      Lai PG, Baughn MR, Yao MG, Walla NK, Elliot VS, Xu Y;
PI      Honchehl CD, Yue H, Ding L, Gietzen KJ, Ison CH, Lu DM;
PI      Hatfield AJA, Ghandi AR, Thangavelu K, Sanjagwala MM, Tang YT;
PI      Ramkumar J, Griffin JA, Swannaker A, Azimzai Y, Saperstein SK;
PI      Burford N, Lee EA, Lu Y, Tran UK, Marquis JP;
XX      WPI: 2002-713453/77.
DR      N-PSDB; AAL49929.
XX
XX      New human molecules for disease detection and treatment (MDPT), useful
PI      for diagnosing, treating and preventing diseases or conditions
PI      associated with the aberrant MDPT expression, e.g. cancer, AIDS,
PI      asthma, diabetes, hepatitis.
XX
XX      Claim 1; Page 128-130; 177p; English.
XX
XX      The present invention relates to human proteins and coding sequences of
XX      molecules for disease detection and treatment MDPT. The sequences can be
XX      used in the treatment of diseases associated with the decreased
XX      expression of overexpression of MDPT, such as cell proliferative (cancer,
XX      atherosclerosis, hepatitis), autoimmune/inflammatory (e.g. AIDS,
XX      allergies, Addison's disease, asthma), developmental (dwarfism, renal
XX      tubular acidosis), neurological (e.g. stroke, Parkinson's disease,
XX      epilepsy) and cardiovascular (congestive heart failure, myocardial
XX      infarction, angina pectoris) disorders. The present sequence is a protein
XX      of the invention.
XX
XX      Sequence 725 AA:
XX
XX      Alignment Scores:
XX      Pred. No.: 0 Length: 725
XX      Score: 3732.50 Matches: 723
XX      Percent Similarity: 98.23% Conservative: 0
XX      Best Local Similarity: 98.23% Mismatches: 2
XX      Query Match: 68.62% Indels: 11
XX      DB: 23 Gaps: 1
XX
XX      US-10-001-857-42 (1-3096) x AAO19400 (1-725)
QY      556  ATGCTTATGAAGCTTCTGTAGATGATGACGATTCAGAGTGGAGCTCAGTATGCCGAA 615
Db      1  MetValMetCysAlaSerValAspAspAspSerGlyTrpGluLeuSerMetProGlu 20
QY      616  AAAATGAGAAAGCAATACAAACTGGGTGACATTCACCAAGATTTGAAGAGCTGTG 675
Db      21  LysMetGluYrSerSerAsnThrAsnTrpValAspIleThrGlnAspPheGluAlaCys 40
QY      676  CGAGATTTAAAGTTGGAGAACTTACTCATGATTAAGCTATTTGCTTTTGAAGCCATG 735
Db      41  ArgGluLeuYrLeuGlyGluLeuLeuHsAspYrLeuPheGlyLeuPheGluAlaMet 60
QY      736  TCTCTCTTGAATATGATGATCCCAAGATGAGTGGCTGATTTGGAACCAAGTAAAT 795
Db      61  SerIleAlaGlnMetCysPheProYrMetAspIleGlyMetIleGlyAsnGlnValAsn 80
QY      796  CGAAAGTTCTCAATTTTGAACAAGCTATGCAAGATGACACTATTAATTAAGATCTC 855
Db      81  ArgYrValLeuAsnPheGlnGlnAlaIleYrAspGlyThrIleYrYrYrYrYrYrYr 100
QY      856  ACCGTGCTGACATGATGAGGATTAAGATACATGATTTGCTGCTTGAATGAGCTGTA 915
Db      101  ThrLeuProGluLeuIleGlyIleMetAspThrCysPheCysLeuIleThrTrpLeu 120

```

QY 916 GAAGGCATTCACTGGACAGACAGATATTACGTCCTTACATTCATTAATCCAGACTT 975
 Db 121 GtGtGtHhIserLeuAlaGlnThrValPheThrCysLeuTyrIleHisAsnProAspPhe 140
 QY 976 ATAGAAATCTCTGCTATGAAGGCTTTTGTCTTGGGAATCTTGAAATCTGTGACATTGCA 1035
 Db 141 HlegLaspProAlaMetLysAlaPheAlaLeuGlyIleLeuLysIleCysAspIleAla 160
 QY 1036 AGGGAATAATTAATTAAGCTGCTTTTGAAGAGAGATTTTACGATCAATGACTTAT 1095
 Db 161 ArgGlnLysValAsnLysAlaAlaValPheGlnGlnLaspPheGlnSerMetThrTyr 180
 QY 1096 GGAATTTAAATGCGCTAACAGTGTGACAGATCTTCAGATTACAGGATGCTAAAGATGTG 1155
 Db 181 GlyPheLysMetAlaAsnSerValThrAspLeuArgValThrGlyMetLeuLysAspVal 200
 QY 1156 GAGATATACATTCGAAAGAAAGTAAGAGTAAAGTACGAAAGTACGAAAGAGAAAGAT 1215
 Db 201 GlnAspAspPheClnArgValLysSerThrArgSerArgGlnGlnLaspPheThr 220
 QY 1216 CCAGAACTGAACTAGACACCAACATATTTAGCATTTTACAGAGTGAATTTACT 1275
 Db 221 ProGlnValGlnLeuGlnHhIsgInGlnCysLeuAlaValPheSerArgValLysPheThr 240
 QY 1276 CGTGTCTTACTGACAGTCTTATAGCTTCTTACTAAGAAAGACCAAGTGTCTGTGAGAA 1335
 Db 241 ArgValLeuLeuThrValIleuIleAlaPheThrLysGlnLysGlnThrSerAlaValAlaGln 260
 QY 1336 GCTCAAAATTTAGTGTTCAGAGCAGCATCTTCTTCCATTCATTAATTAATTCAT 1395
 Db 261 AlaglnLysLeuMetValGlnAlaAlaAspLeuLeuSerAlaIleHisAsnSerLeuHis 280
 QY 1396 CATGCACTCCAGGCCCAATGATCTACMAAAGAGATCATCCATTAATGAGGGTTT 1455
 Db 281 HhIsgLysLeuGlnAlaGlnAsnAspThrThrLysGlnLysAspIleMetGlyPhe 300
 QY 1456 GAACCCCTTGGAAACGAGGCTACTCCACCTACCTCCCTGATGCAATTAATTAATTAAT 1515
 Db 301 GlnProLeuValAsnGlnAlaGlnLeuLeuProThrPheProArgTyrAlaLysIleIle 320
 QY 1516 AAAAGGAAAGATGTGAACTATTTTGCAGATTAATAGATTAATTAATTAATTAAT 1575
 Db 321 LysArgGlnGlnMetValAsnTyrPheAlaArgLeuLeuAspArgIleLysThrValCys 340
 QY 1576 GAGGTGTGATTTAAACAAATTTACATGATCTGATCTTCTGTAATTTAGTGA 1635
 Db 341 GlnValValAsnLeuThrAsnLeuHisCysIleLeuAspPheCysGlnPheSerGln 360
 QY 1636 CAGTCACCATGTGCTCTTCAAGATCTGTGTACAAACCACTTCTGTGTGATTAACAA 1695
 Db 361 GlnSerProCysValLeuSerArgSerLeuLeuGlnThrPheLeuValAspAsnLys 380
 QY 1696 AAGGCTTTGGAAGTCACTCTCATGCAAGACATGGTGAAGATGACTGGCTCTTTTGTG 1755
 Db 381 LysValPheGlnTyrHhIsgLeuMetGlnAspMetValLysAspAlaLeuArgSerPheVal 400
 QY 1756 AGATCTCTCGAGTCTTCCCAAGTGTACTATTAATTAATTAATTAATTAATTAATTAAT 1815
 Db 401 -SerProPro-ValLeuSerProLysCysTyrLeuTyrAsnAsnHhIsgAlaLysAspC 420
 QY 1816 GTATCGACTCTCTTGTATCACTGCTGTGGCATCTGTAGCTTATCAATCAATG 1875
 Db 420 YstIleAspSerPheValThrHisCysValAlaGlnProPheCysSerLeuIleGlnHhIsg 440
 QY 1876 GACATAACAGGCTCGACAGAGATTAAGCTGTGCTCATATTTCTTGAGAAATTTGCCACT 1935
 Db 440 LysHisAsnArgAlaArgGlnArgAspLysLeuGlnHisIleLeuGlnGlnPheAlaThrL 460
 QY 1936 TGCAGATGAGTTATGACATTTATTTAATAGCAGAGAGATTAATGACAGGCTTCA 1995
 Db 460 eugLasp-----Gln-AlaGlnLysValAspAlaAlaLeuHhI 472
 QY 1996 CACCATGCTGTGAAAGAGAAACCCAAAGCAACATTTGGCTGTAGTACTGGGT 2055

Db 472 sThrMetLeuLeuLysGlnGlnProGlnArgGlnHhIleuAlaCysLeuGlnGlyThrTyrVal 492
 QY 2056 CCTTTACATTAACCTTGGCATTTATGATPACACTTCTTAGTGCTTGAATTTGGAAT 2115
 Db 492 HleuTyrHisAsnLeuAlaGlnIleMetIleGlnTyrLeuLeuSerGlyPheGlnLeuGlnLe 512
 QY 2116 CTACAGATGACAGGACTATTAATATATTTGGATATCTCTGTAATTCCTTTAGCATG 2175
 Db 512 uTyrSerMetHhIsgLysTyrTyrTyrIleTyrIleTyrTyrLeuSerGlnPheLeuTyrAlaTr 532
 QY 2176 GTTGATGTCAACATTTGAGTCTGCGGATGCTCTTCAAAATGGCAGAGAAAGATTAATGCA 2235
 Db 532 PheMetSerThrLeuSerArgAlaAspGlySerGlnMetAlaGlnGlnLaspGlyIleMetGln 552
 QY 2236 AGAGCAGAGAAAGCCGTAGTAAAGAAAAAAGAAAAAAGAAAAAAGTTCGCCCAT 2295
 Db 552 uGlnGlnGlnLysGlnArgSerSerLysLysLysLysLysLysLysLysLysLysLysLys 572
 QY 2296 GAGCCGAGATCACAATGAGCCAAAGCATATCAGAACATGTGTGCTGGAATGTTAAAC 2355
 Db 572 uSerArgGlnIleThrMetSerGlnAlaTyrGlnAsnMetCysAlaGlyMetPheLysThr 592
 QY 2356 CATGGTAGATTTGACATGACCGCAAGTACGTAAACGAAAGTTGAGCTTGAATGATGCA 2415
 Db 592 HMetValAlaPheAspMetAspGlyLysValAlaArgLysProLysPheGlnLeuAspSerGln 612
 QY 2416 ACAAGTTCGGATGTAACAACAGTTTGTCTTCATTAACAGTGTGATGACCCGCCAGT 2475
 Db 612 uGlnValAlaArgTyrGlnHhIAspPheAlaProPheAsnSerValMetThrProProVal 632
 QY 2476 GCACTACTTACAGTTCAGAGAAATGTGACCTGATTAATTAATTAATTAATTAATTAAT 2535
 Db 632 HhIsgTyrLeuGlnPheLysGlnMetSerAspLeuAsnLysTyrSerProProGlnLys 652
 QY 2536 TCCTGAATGTATGAGCAGCTAGTAAACATTTTCAACAGGCAAAATGATATTGAAAA 2595
 Db 652 rProGlnLeuTyrValAlaAlaSerLysHhIsgGlnAlaLysMetIleLeuGlnAs 672
 QY 2596 TATTCCTAACCCGACCATGAGGTTAATGAATTTTAAAGTGGCAAAACCAACTTTGT 2655
 Db 672 HhIsgProAsnProAspHisGlnValAsnArgIleLeuLysValAlaLysProAsnPheVal 692
 QY 2656 GGTATGAATTTATGAGCAGAGACACAAAGAAATCTAAAGTTCCTCTGAAATTTGA 2715
 Db 692 IValMetLysLeuLeuAlaGlyGlyHisLysLysGlnLysSerValProProGlnPheAs 712
 QY 2716 TTTCTGTGCTCATTAATATTTTCTGTTGTGAACTTGT 2755
 Db 712 PheSerAlaHisLysTyrPheProValValLysLeuVal 725
 RESULT 3
 ABB60388
 ID ABB60388 standard; Protein; 784 AA.
 AC ABB60388;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 7956.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN MO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.

```

PR      11-JUL-2000; 2000US-0614150.
PA      (PEKE ) PE CORP NY.
PL      Venter JC, Adams M, Li PWD, Myers EW;
XX      WPI; 2001-655860/75.
DR      N-PSDB; ABL04491.
XX      XX
PT      New isolated nucleic acid detection reagent for detecting 1000 or more
PT      genes from Drosophila and for elucidating cell signalling and cell-cell
PT      interactions -
XX      XX
SS      Disclosure; SEQ ID NO 7956; 21bp + Sequence Listing; English.
XX      XX
CC      The invention relates to an isolated nucleic acid detection reagent
CC      capable of detecting 1000 or more genes from Drosophila. The invention is
CC      useful in developmental biology and in elucidating cell signalling and
CC      cell-cell interactions in higher eukaryotes for the development of
CC      insecticides, therapeutics and pharmaceutical drugs. The invention
CC      discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC      sequences (ABL01840-ABL16175) and the encoded proteins
CC      (AAB57737-AAB72072).
CC      CC
CC      The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pat_sequences.
XX      XX
SQ      Sequence       784 AA;

Alignment Scores:
Pred. No.:           2.07e-114          Length:         784
Percent Similarity:   1298.50            Matches:        306
Best Local Similarity: 55.42%             Conservative:   139
Query Match:         23.87%              Mismatches:    265
                                         Indels:        93
                                         Gaps:         18

US-10-001-857-42 (1-3096) x ABB60388 (1-784)
QY      472 GCCGGCGGCGGCCGAGCGGCCTGTGA-----TTTCGGTGGTCCGACAGTGCGTGGCG 525E
Db      34 AlAlAlAlAlAlArghAlaserSerLeugInuInAsglutItpser----- 48
QY      526 GCCGGGGTGAACCACGGGAGAAGTAGGCATAAAG-----GTTATG 564ATG
Db      49 -----ValtlygLucysgLyPheLuasppProAspvalGlnArgThmeLarg 64
QY      565 AAAAGCTTCTGTAGATGATGACGATTCCAGATGGAGTGCAGTACGTCCAGAAAATAATGAG 624ATG
Db      65 SerGIySerAlaInlaelnuIspllethrIntyrPromethis----- 78
QY      625 AAAAGCAATACAACCTGGGTGGACATTAACCCAAGATTTGAAAGACTGTTCGGAATTA 684ATG
Db      79 -----GLyrrPalaspvalThrIrysGluPrhen;sAspAlaCySaIagLUleu 94
QY      685 AAGTTGGGAACTACTTCATGATTAAGCATTTGGCTTTTGAAGCAGTCTGCTATT 744ATG
Db      95 GlmProGlYgIUleuNlaaglInsphetleuPrehgIyleuPrehgIulaImetSerAlalle 114
QY      745 GAAATATGATGCCCAAGTGTGATGCTGGCATGATTGSAAACCAAGTTAATCGAAAAATT 804ATG
Db      115 GluIleMetAspProIysmetAspValgIyMetGlyPheAspIrySGlInAspleuProPro 134
QY      805 CTCATTTTGAACAAGCTATCAAGAGTGAGCATTAATAAATTAAGATCTCACCTTGCT 864ATG
Db      135 ProSerPhcGUlaAlaAlaIleAlahrhrgIalAleIylSleuAspAspleuthrProSer 154
QY      865 GAACTGATAGGAGATTATGCATGCATGTTTGTGCTGTTTGATTAACGTGATTGAAGGCCAT 924ATG
Db      155 GUleuIIleGIyIleTyraSpAlaleuBheserCYleuValserTrpleuGIuGIyasn 174
QY      925 TCACGTGCACAGACAGIATTTACGGTCCCTTACATTCCTATTCAGACTTATTAAGAAT 984ATG

```

Db	175	SerMetAspGlnValIleuPheThrCysLeuTyrlleuHISalAlProAlaGlnIleLeuAsp	194
QY	965	CGTGTATGAAGGCTTTTGCTCTGGGAATCTTTGAAATCTGTGTGATTCGACAGGGAAGAAA	1044
Db	195	LysAlaLeuAlaArgValPheCysThrAlaValArgMetIleValIleLeuLysIle	214
QY	1045	GTAATATAAGCTGCTGTTTTTGAAGAGGAAGATTTTCAGTCAATGACTTATGCA-----	1098
Db	215	IleAlaValAlaAlaValAlaSerGlnGluAspPheGln-----LeuTyrlGlyAsnSer	232
QY	1099	---TTTAAATAGGCTTAACAGTGTACAGATCTTGAGATCTTACAGCAAGCTGAAAAGATGTG	1155
Db	223	AlaIleuLeuAlaIlaGluLysAlaGlnProAlaThrValTyrlSerSerLeuLysAspVal	252
QY	1156	GAGATGACATGCAGAAAGAGATGAAGATCTCGAATGCGACAGAGAAAGAAAGAT	1215
Db	253	GluAspGlnIleuIleArgLysCysValLysLeuThrSerThr-----	266
QY	1216	CCAGAGTTGAACATAGAACCAGCAATGTTTTACAGATATTTCAGCAGAGTGAATTACT	1275
Db	267	-----GluAspTrpMetAlaValAlaHISArgLeuAspPheMet	279
QY	1276	CGTGTGTACTGACAGTGGCTTTATAGCCTTTACTAGAAAGAGACAGCTGCTGTTCAGAA	1335
Db	280	ArgHISLeuPheGlnValIleTyrlHISValGlnIleMetAlaSerAsnAspThrValAsp	299
QY	1336	GCTGCA-----AAATGATAGCTTCAGACAGCAAGATCTTCTTCTGCCATTGAT	1383
Db	300	AspLysValAspIleTyrlLysIleLeuLeuValAlaSerGlnMetLeuProGlyIleArg	319
QY	1384	AATTCATTGTCATCGCATCGCATCCAGGCCAGAAATGATCTCAAAAGAGA---GATCATGCA	1440
Db	320	AsnThrLeuAspArgGlyThrGlnProGlu-----LysGlySerAspAlaPro	335
QY	1441	ATTATGATGGGTTTGGAAACCCCTTTGTAAACAGGCTACTTCCACCTTACCTTCCCTCGA	1500
Db	336	AsnProMetGlyPheSerProAlaGlnIleAspArgSerGlnProAlaPheProArg	355
QY	1501	TATCAAAATAATATTAAAGGAAAGAAATGTGACATTTTGCAAGATTAATAGATAGA	1560
Db	356	SerIleLysIleArgAspArgProSerSerTyrlGlnPheLeuGlnIleMetIleSerArg	375
QY	1561	ATAAAAACGTGCTGTGAGGTGTGGAATTAAACAATTTACATTTGATCTCGATTTTTC	1620
Db	376	PheLysTyrlAlaCysLysValIleThrLysTyrlLysAspTyrlTyrlSerAlaLeuAsnPhe	395
QY	1621	TGTAATTTTGTGAACAGTCA---CCATGTCCTTTTCAAGATCTCGTTACAAACCACT	1677
Db	396	IleGluTyrlSerLysLysSerGlyGlnIleCysIleLeuSerTrgSerValLeuGlnThrLeu	415
QY	1678	TTCTCGTGTGATGAACAAAAGGCTCTTTGGAACATCTTCAGAACAGCATGTGAAAGAT	1737
Db	416	PheSerAlaAsnMetArgMetAlaHISGlyLysLeuProMetLysGlnPheLeuArgHIS	435
QY	1738	GCACTTGGGCTTTTGTGCAATCTCTCGAGTCTTTCGCCAAGTGTACTCTATATATA	1797
Db	436	SerValAlaIleValPheAsnSerProProValLeuAsnAlaValSHSProValAlaAla	455
QY	1798	ATCCACAGGCTTAAGAACTGTATGACATCCCTTTTATCTACTGTGT---CGGCGAT	1851
Db	455	SProlLysValGlnGlnIleHisLeuGlnIleAsnPheHISGlyCysIleAsnMetAsnThrP	475
QY	1852	TTCGTAGCTTATTTCAGATCSATGACATTAACAGGCTGTGACAGAGATTAAGCTTGCTG	1911
Db	475	heThrGlnPheIleArgGlyLeuCysGlyPheAsnArgAlaArgGlnArgAspLysLeuAla	495
QY	1912	ATATCTTTGAGAAATTTGGCACTTGCAG---GATGAGTTATAGA	1953
Db	495	rgLeuIleGluAsnPheAspThrIleGlnIleValAspAlaAlaArgLysAspSerMetMet	515
QY	1954	CATTTATTTTAAATAGCAGA-----GAAGCTGTAGCAGCGCTTCAACACCATGCTGT	2007
Db	515	snGlnIleuAlaAsnGluArgAlaMetGlnGlyArgGlnLysProMetAlaThrAlaIleu-	533

SQ Sequence 239 AA:

Alignment Scores:

Pred. No.:	7,99e-109	Length:	239
Score:	1237.00	Matches:	238
Percent Similarity:	99.58%	Conservative:	0
Best Local Similarity:	99.58%	Mismatches:	1
Query Match:	22.74%	Indels:	0
DB:	23	Gaps:	0

US-10-001-857-42 (1-3096) x ABG69816 (1-239)

QY 475 GCGGCGCGCGAGCGCGCTGCTGATATTCGCTGCTCGGACAGTGCCTGGCGCGCGCTG 534
 1 AAlAlAlAGluAlAlAlSerLeuPheProTyrSerGlyGlyCysValAlAlAlVal 20
 QY 535 ACCACGGGGAATAGCATTAATGCTTAAGGCTTCCTGATATATACGATTAGGA 594
 21 ThrThrGlyGluValGlyIleMetValMetLysAlaSerValAspAspAspSerGly 40
 QY 595 TGGAGCTAGTATGCGCAGAAAAAATGAGAAAAAGCATATCAAACTGGGTGACATTACC 654
 41 TrpGluLeuSerMetProGluLysMetGluLysSerAsnThrAsnTrpAlaSplIleThr 60
 QY 655 CAAGATTGGAAGACCTGTCTCGAGATTAAGTTGGGGAAGTCTTATGATTAAGCTA 714
 61 GluAspPheGluGluValCysArgGluLeuLysLeuGlyGluLeuLeuHisAspLysLeu 80
 QY 715 TTGGCTCTTTTGAAGCCATGCTGCTGATTAAGATGATGATGATGATGATGATGATG 774
 81 PheGlyLeuPheGluValMetSerAlaIleGluMetMetAspProLysMetAspAlaGly 100
 QY 775 ATGATTGGAACCAAGTAAATCGAAAAAGTTCTCAATTTTGAACAAGCTATCAAGATGCG 834
 101 MetIleGlyAsnIleValAsnArgLysValLeuAsnPheGluGluAlaIleLysAspGly 120
 QY 835 ACATTAATAATTAACATCTCACCTGCTGGAAGTGGGATTAAGTATGATCAATGTTT 894
 121 ThrIleLysIleLysAspLeuThrLeuProGluLeuIleGlyIleMetAspThrCysPhe 140
 QY 895 TGCTGTTTGAATACTGTGTTAAGGCCATTCATCTGCGACAGACAGATATTTCAGTGCCTT 954
 141 CysCysLeuIleThrTrpLeuGluGlyHisSerLeuValGluThrValPheThrCysLeu 160
 QY 955 TACATTCATTAATCCAGACTTTATAGAAATCCTGCTGCTAAGGCTTTTCTTGGGAATC 1014
 161 TyrIleHisAsnProAspPheIleGluAspProAlaMetLysAlaPheAlaLeuGlyIle 180
 QY 1015 TTGAAAATCTGTGACATTCAGAGGAAAAAGTAATAAGTCTGTTTGAAGAGGA 1074
 181 LeuLysIleCysAspIleAlaArgGluLysValAsnLysAlaAlaValPheGluGluGlu 200
 QY 1075 GATTTTCATCAATGCTTATGATTTTAAATGGCTTAACAGTGTGACAGATTTTGAGCTT 1134
 201 AspPheGlnSerMetThrTrpGlyPheLysMetAlaAsnSerValThrAspLeuArgVal 220
 QY 1135 ACAGGCGCTCTAAAGATGTGAGGATGACATGCAAGAAGAGTAAAGATCTCTGCA 1191
 221 ThrGlyMetLeuLysAspValGlyAspAspMetGlnArgArgValLysSerThrArg 239

RESULT 5

AAU32414

ID AU32414 standard; Protein, 108 AA.

XX AU32414;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #2905.

XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukemia.

XX OS Homo sapiens.
 XX EN WO200179449-A2.
 XX PD 25-OCT-2001.
 XX PF 16-APR-2001; 2001WO-US08656.
 XX PR 18-APR-2000; 2000US-0552929.
 XX PR 26-JAN-2001; 2001US-070160.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Dermanac RT;
 XX WPI; 2001-611725/70.
 XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 PS Claim 20; Page 608; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation, to regulate haematopoiesis, and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukemias. AU29510-AU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

SQ Sequence 108 AA;

Alignment Scores:

Pred. No.:	2.28e-28	Length:	108
Score:	392.00	Matches:	80
Percent Similarity:	96.39%	Conservative:	0
Best Local Similarity:	96.39%	Mismatches:	2
Query Match:	7.13%	Indels:	2
DB:	22	Gaps:	0

US-10-001-857-42 (1-3096) x AU32414 (1-108)

QY 2944 ACCACTTAAGAAGCAGGTTTTCATTCATGATTAAGGCTGCAAAAGTATTC 2885
 1 ThrThrLeuLysGlnGlnPheSerPheMetSerTrpLysAlaValLysLeuValPhe 20
 QY 2884 CTATTAATGAGTTTTCACCAAGAAATCCATCTCACTCCCGTGAAGTGAAGCTGTG 2826
 21 LeuIleMetSerCysTrpProArgAsnProSerHisPhePro***CysGlyAlaTrpVal 40
 QY 2825 ATGTGCCCTTAAAGATGGGTGGAAGAAGATCTGCGCCCTTATGGCCACCTCCCA 2766
 41 MetCysProLeuAlaGlyValGlySerGluArgLeuLysProPheMetAlaThrSer-Gl 60
 QY 2765 GTCTCTCTCAAAACAGTTTCAACACAGAAAAATTTATAGACAGAGAAATCAATTGAG 2706
 60 nSerLeuSerAsnLysPheHisAsnArgLysIlePheMetSerArgGluIleLysPheAr 80
 QY 2705 GAGGAC 2659
 80 gAggAsn 82

RESULT 6

AB93650 standard; Protein; 681 AA.

ID ABB93650

AC ABB93650

31-MAY-2002 (first entry)

Herbicide; active polypeptide SEQ ID NO 2861.

Herbicide; plant; agriculture; herbicide.

Arabidopsis thaliana.

WO200210210-A2.

07-FEB-2002.

28-AUG-2001; 2001WO-EP09892.

28-AUG-2001; 2001WO-EP09892.

(FARB) BAYER AG.

Tietjen K, Weidler M;

WPI; 2002-269010/31.

Identifying plant target proteins for herbicidally active compounds,

comprising aligning and comparing nucleic acid or amino acid sequences

from plant with nucleic acid or amino acid sequences from non-plant

organisms -

Claim 5; SEQ ID NO 2861; 261pp + Sequence Listing; English.

The invention relates to identifying target proteins

(ABB90790-ABB94016) for herbicidally active compounds, comprising

aligning and comparing nucleic acid or amino acid sequences from plant

with nucleic acid or amino acid sequences from non-plant organisms using

suitable search parameters, where plant sequences having an E-value

greater by a factor of 3 than the E-value of most similar non-plant

sequences are selected. The polypeptides or nucleic acids encoding them

are useful for identifying modulators. The identified modulators are

useful as herbicides.

Sequence 681 AA;

Alignment Scores:

Pred. No.: 1.2e-05 Length: 681

Score: 158.50 Matches: 52

Percent Similarity: 38.65% Conservative: 11

Best Local Similarity: 31.90% Mismatches: 60

Query Match: 2.88% Gaps: 40

US-10-001-857-42 (1-3096) x ABB93650 (1-681)

QY 608 ATACTGAGCTCCATCTGATGTCATCTACAGAGCTTTATTAACATTAATGCTT 549

Db 60 ValValSerSerProProSerSerSerProPro-----Pro 72

QY 548 ACTTCTCCGTCGTCACCGCGCGCGACGACGACTGTCGG----- 509

Db 73 SerProProValIleThr-SerProProProThrValAlaSerSerProProProVa 92

QY 508 -----ACCAAGGAATACGACGCGCGCTCGCGCGCGCG----- 476

Db 92 ValIleAlaSerProProProSerThrProAlaThrThrProProAlaProProGlnTh 112

QY 475 -----CGCGCGCGCGCTTCACGCGCGCGCGCGCGCGCGCGCGCGT 423

Db 112 ValSerProProProProProAspAlaSerProSerProProAlaProThrThrThr 132

QY 422 GCCCGGAGACCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 363

Db 132 nPro-----ProProLysProSerProSerProGluThrProSerProPr 149

QY 362 CACACCG 303

Db 149 OGlyGluThrProSerProProLysPro-----SerProSerThrProTh 164

QY 302 GAGATTAGAGACGATCGCGACCGCGAGATTATCTTTTACCGCTGTGAGTCTG 243

Db 164 rProThrThrThrThrSerProPro-----ProProProAlaThrSe 178

QY 242 GGTAGTGGCGCGTTCGCGGAAAGAGCGCGACCGCGAGTCTAGAGCGCGCGCTG 183

Db 178 rAlaSerProProSerSerSerProProThrAspProSerThrLeuAlaProProThrPr 198

QY 182 GCTGCGC 176

Db 198 CleuPro 200

RESULT 7

ABG14149 standard; Protein; 156 AA.

ID ABG14149

AC ABG14149

18-FEB-2002 (first entry)

Novel human diagnostic protein #14140.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dymanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS78336.

New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

biodiversity -

Claim 20; SEQ ID No 44508; 103pp; English.

The invention relates to isolated polynucleotide (I) and

polypeptide (II) sequences. (I) is useful as hybridization probes,

polymerase chain reaction (PCR) primers, oligomers, and for chromosome

and gene mapping, and in recombinant production of (II). The

polynucleotides are also used in diagnostics as expressed sequence tags

for identifying expressed genes. (I) is useful in gene therapy techniques

(II). (II) is useful for generating antibodies against it, detecting or

quantitating a polypeptide in tissue, as molecular weight markers and as

a food supplement. (II) and its binding partners are useful in medical

imaging of sites expressing (II). (I) and (II) are useful for treating

disorders involving aberrant protein expression or biological activity.

The polypeptide and polynucleotide sequences have applications in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG0377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pct_sequences.

XX Sequence 156 AA:

Alignment Scores:

Pred. No.:	1.61e-05	Length:	156
Score:	153.50	Matches:	52
Percent Similarity:	34.66%	Conservative:	9
Best Local Similarity:	29.55%	Mismatches:	56
Query Match:	2.79%	Indels:	59
DB:	22	Gaps:	7

US-10-001-857-42 (1-3096) x ABG14149 (1-156)

```

QY 556 TTATGCTTACTTCTCCGCTGTCACCCGCGCCGACCTGTCGAGACCGGAAATA 497
   |||||
Db 1 LeuCySerLeuSerSerLeuProProProPro----- 12
QY 496 ACGAGCGCGCTCGCGCGCGCGCGCGCGCTCTCAGCCGACCGCGAGCGT 437
   |||||
Db 13 ---ProProProProProProProProPro----- 22
QY 436 GCACGCATCGCTATGCGCGGACGCGGACCGCGCGCTCTCTACTCCGCGGACT 377
   |||||
Db 23 ---ProProProProProProProProProProProPro----- 37
QY 376 TCGCGCTCCCGCGCACACCCCTCCCTGCGACCGCTCCGCTC-----GCGCGA 326
   |||||
Db 38 ---SerProArgSerProProProProProProProProProProProPro----- 55
QY 335 TCGCGCGCGCTCTCTCTCTCTCGGAGATAGACGATCGGAGACCGGAAATTATCTT 266
   |||||
Db 56 SerProProSerGluLeuGlySerSerAlaGluProTyrPleuArgProGlyThrTyrVal 75
QY 265 TTTCACCT-----CTGCGAGTCTGCGTGGTAGTGGCGGTTCCCGAAAGAGCG 215
   |||||
Db 76 ***ProProProLeuThrPheSerGlnGlnGlnAlaAlaSerSerAlaGlnLys***Asn 95
QY 214 GAGCGCGAGTCTCA---GAGCGCGCGCGCTGCGCTGCGCTCGGGAATTTCTTAA 158
   |||||
Db 96 CysProGlyGlyLeuSerProGlnProValProValThrValPro----- 109
QY 157 AGGTACCGGAGGCGGAAAGAACCTGTGAGAGTTGCGTAGACCTTAGTGTGCGAGC 98
   |||||
Db 109 ----- 109
QY 97 ACGAAGAGAGCGCAAGCTCGAGAGCAACAGCTGACTAGTGGCT 50
   |||||
Db 110 ---GluArgGlnProAspAlaAlaProXnHisSerLeuLeuPro 124

```

RESULT 8

ABG1504

ID ABB91504 standard; Protein; 708 AA.

AC ABB91504;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 715.

KX Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

XX WO200210210-A2.

PD 07-FEB-2002.

XX

PF 28-AUG-2001; 2001WO-EP09892.

XX 28-AUG-2001; 2001WO-EP09892.

XX (FARB) BAYER AG.

PI Tietjen K, Weidner M;

DR WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,

PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant

XX organisms -

XX Claim 5; SEQ ID NO 715; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.

XX Sequence 708 AA:

Alignment Scores:	4.11e-05	Length:	708
Pred. No.:	153.00	Matches:	59
Score:	153.00	Conservative:	17
Percent Similarity:	39.38%	Mismatches:	71
Best Local Similarity:	30.57%	Indels:	46
Query Match:	2.78%	Gaps:	8

US-10-001-857-42 (1-3096) x ABB91504 (1-708)

```

QY 608 ATACTAGCTCCGATCTGATCATCTTACAGAAAGCTTTCATACCATATGCT 549
   |||||
Db 101 ValIleProSerProProProSerAlaSerProProProAlaLeuValProProLeuPro 120
QY 548 ACTTCTCC-----GTGTCACCGCGCGCGCGACG 519
   |||||
Db 121 SerSerProProProProAlaSerValProProProArgPro-SerProSerProProI 140
QY 518 CACTGTCGCGAGC-----ACGAAATAGAGAGCGCGCTCGCGCGCG-- 476
   |||||
Db 140 eleuValArgSerProProProSerValArgProIleGlnSerProProProProProse 160
QY 475 -----CGGCGCGCTCCCTCTGAGCCAGCCGAGCGAGCGGTGC 435
   |||||
Db 160 rAspArgProThrGlnSerProProProProSerProProSerProProSerGluArgPr 180
QY 434 ACGCATG-----CGTATGCCCGGAGCGCGAGCCCGCCCTC 396
   |||||
Db 180 oThrGlnSerProProSerProProSerGluArgProThrGlnSerProProProProse 200
QY 395 TCCCTTACTCCCGAGACTTTCGCTCCCGGCGCACACCCCTCCCTCGCACCGCTCC 336
   |||||
Db 200 rProProSerProProSerAspArgProSerGlnSerProPro-----ProProProG 218
QY 335 GTTCGCGCGATGCGCGCGCTCTCTCTCGGAGATTAGAGAGATGCGAGACCGGA 276
   |||||
Db 218 uAspThrLysProGlnProProArg-----ArgSerProAsnSerProPr 233
QY 275 AGTTATCTTTTTCACCTCTGCGAGTCTGGTAGTGGGCGGCTCCCGGAAAGAGCG 216
   |||||
Db 233 oProThrPheSerSerPro-ProArgSerProProGluIleuValProGlySerAsn 253
QY 215 GAGCGCGAGTCTAGAGCCCGCGCGCTGCGGTGCGCTCTGGAATTCTTAGAAG 156
   |||||

```

DB 253 sm-----ProserGlnAsnAsnProThrLeuAyrProProLeu----- 265

QY 155 GTGACCGGAGAGCGGAGAAAGACCTGTGAGCTTG 121

DB 266 --AspAlaProAsnSerThrAsnAsnSerGlyTle 276

RESULT 9

ABUS2794

ID ABUS2794 standard; Protein; 120 AA.

AC ABUS2794;

XX

DT 14-APR-2003 (first entry)

XX

DE Human transmembrane protein DKFzphb2_82g14 homologue #2.

XX

KW Human; gene therapy; vaccine; disease treatment; detection.

XX

OS Homo sapiens.

XX

PN WO200112659-A2.

XX

PD 22-FEB-2001.

XX

PF 18-AUG-2000; 2000WO-IB01496.

XX

PR 18-AUG-1999; 99US-0149499.

XX

PR 28-SEP-1999; 99US-0156503.

XX

PA (GEHU-) GERMAN HUMAN GENOME PROJECT.

XX

PI Wiemann S;

XX

DR WPI; 2001-327840/34.

XX

PT Nucleic acids having the sequences of clones isolated from libraries of

XX

PS different human tissues, useful in recombinant DNA methodologies -

XX

XX Example III; Page 359; 1095pp; English.

XX

CC This invention describes novel polynucleotides and polypeptides isolated

CC from human cDNA libraries which can be used for gene therapy or in

CC vaccines. The polynucleotides of the invention and antibodies encoded by

CC them may be used in the prevention, diagnosis and treatment of diseases

CC associated with inappropriate polypeptide expression. The products of the

CC invention may also be used to identify modulators of expression and

CC activity and to down regulate expression and activity. The antibodies of

CC the invention may also be used as diagnostic agents for detecting the

CC presence of polypeptides in samples. This sequence represents a homologue

CC of a polypeptide described in the disclosure of the invention.

XX

SQ Sequence 120 AA;

Alignment Scores:

Pred. No.: 6.5e-05 Length: 120

Score: 146.50 Matches: 44

Percent Similarity: 38.93% Conservative: 7

Best Local Similarity: 33.59% Mismatches: 53

Query Match: 2.67% Indels: 27

DB: 22 Gaps: 5

US-10-001-857-42 (1-3096) x ABUS2794 (1-120)

QY 541 CCGTGTGTCACCGCGCGCCACGACACTGTCCGACACGAGAAATPACGACCGCCCT--- 485

DB 6 ProProGlyProProProProGlyGlnValLeuProProProLeuAaGlyProProAsn 25

QY 484 -----CGGCGCGCGCGCGCCCTCTCAGCGCCGCGCGACCGCGACGCTGCA 434

DB 26 ArgGlyAspAlaGlyProProProProValLeuPheProGlyGlnProPheGlyGlnProPro 45

QY 433 CGCATGCGTATGCGCGCGGACCGCGACCGCCCGCCCTCTACTCCCGGGAATTG 374

DB 46 LeuGlyProLeuPro-----ProGlyProProProProValProGlyTyrGly 61

QY 373 CGCTCCCGCGGACACACCCCTCC-----TCGCGACGCGCTCGGCGCATG--- 323

DB 62 ProProProGlyProProProProGlnGlnGlyProProProProProGlyProPhePro 81

QY 322 --CGGCGCGCTTCTCTTCTCGCGAGATTAGAGACATCGCGAGACCGGAATTATCTTT 266

DB 82 ProArgProProGlyProLeuGlyProProLeuThrLeuAlaProProHisLeuPro 101

QY 265 TTTCACCTCTGTGGAATCCTGGGTAGTGGCGGTTCCCGAAAGAGCGGACCGGGA 206

DB 102 GlyProPro-----ProGly 106

QY 205 GTCTCAGACCGCGCGCGCTGTGCGCTGCGCTCT 173

DB 107 AlaProProProAlaProHisValAsnProAla 117

RESULT 10

AAM52322

ID AAM52322 standard; Protein; 574 AA.

XX

AC AAM52322;

XX

DT 18-JAN-2002 (first entry)

XX

DE WASP homolog protein.

XX

KW Actin polymerisation; Ena/VASP; vasodilator-stimulated phosphoprotein;

XX

KW metastatic cancer; parasitic infection; cytotoxic; WASP.

XX

OS Schizosaccharomyces pombe.

XX

PN WO200171356-A2.

XX

PD 27-SEP-2001.

XX

PF 21-MAR-2001; 2001WO-FR00843.

XX

PR 22-MAR-2000; 2000FR-0003637.

XX

PA (CNRS) CENT NAT RECH SCI.

XX

PA (CURT-) INST CURIE.

XX

PI Fradelizi J, Friederich E, Golsteyn RM, Louvard Y, Noireaux V;

XX

DR WPI; 2001-639148/73.

XX

PT Identifying modulators of actin polymerization, potentially useful for

XX

PT treating tumor metastasis and parasitic infection, using proteins that

XX

PT contain Ena/VASP binding sites -

XX

XX Claim 13; Pages 107-109; 109pp; French.

XX

CC The present invention relates to a method for identifying modulators of

CC actin polymerisation. The method involves using proteins that contain at

CC least one binding motif for proteins of the Ena/VASP

CC (vasodilator-stimulated phosphoprotein) family in the preparation of

CC reagents for identification/screening of molecules that modulate

CC formation of the actin cytoskeleton. The proteins used in the method

CC (i.e. the proteins with binding motif(s) for Ena/VASP proteins) do not

CC bind to the Arp2/3 protein complex. The modulators identified by the

CC method are potentially useful for treating disorders of actin

CC polymerisation, e.g. metastatic cancer or parasitic infection; and as

CC cytotoxic agents. The present sequence one such protein with binding

CC motif(s) for Ena/VASP proteins, which was used in the method of the

XX

SQ Sequence 574 AA;

Alignment Scores:

Pred. No.: 0.000153 Length: 574

Score: 146.50 Matches: 53
 Percent Similarity: 39.41% Conservative: 14
 Best Local Similarity: 31.18% Mismatches: 64
 Query Match: 2.67% Indels: 39
 Gaps: 8

US-10-001-857-42 (1-3096) x AAM52322 (1-574)

```

QY 544 CTCCTGTCACCCGCGCCGACCTGTCGAGACCGAATAAGACCGCCT 485
   |||||
Db 336 LeuProProProProProProProProProSerAlaAlaGlySerLeuProProPro 355
QY 484 CGG-----CCGCGCGCGCGCGCTCCCTCTCAGCCGCGACCGGACGCT 437
   |||||
Db 356 GlnGlyArgSerAlaProProProProProProProSerAla---ProSerThylArg 374
QY 436 GCACGATGCGTATGCGCGGACCGC-----GACCGCGCGCGCTCCCTCTA 389
   |||||
Db 375 GlnProProProProProProProProProSerAlaAlaValSerAsnProProAla 394
QY 388 CTCCTGCGGACCTTCGCGCTCCCGCGCACACC----- 356
   |||||
Db 395 IleProGlyArgSer---AlaProAlaLeuProProLeuGlyAsnAlaSerArgThSer 413
QY 355 -----CTCCCTGCGCGCGCGCTCCCTCTCGCGCGCA 326
   |||||
Db 414 ThrProProAlaProThrProProProSerLeuProProSerAlaProProSerLeuProPro 433
QY 325 TGGCGCGCGCTCTCTCTCTCGGAGATTAGACAGCATCGGACCGGAGTTATCTT 266
   |||||
Db 434 SerAlaProProSerLeu-----PrometGlyAlaProAlaAlaProProLeuProPro 451
QY 265 TTTCACCTCTGCGGAGTCTGCGTATGCGC-----GTTCC 227
   |||||
Db 452 SerAlaProProLeuProProProAlaGlyMetProAlaAlaProProLeuProPro 471
QY 226 CGGAAAGAGCGAGCGCGCGAGTCTCAGACCGCGCGCTGCGCTGCTGGAGAT 167
   |||||
Db 472 AlaAlaProAlaProProProAlaProAlaProAlaProAlaAla---ProValAlaSer 490
QY 166 TTCTTAGAAAGGTGACGCGGACCGGAGG 137
   |||||
Db 491 IleAlaGluLeuProGlnGlnAspGlyArg 500

```

RESULT 11

AAG67370
 ID AAG67370 standard; Protein; 574 AA.

AC AAG67370;

DT 13-NOV-2001 (first entry)

DE Amino acid sequence of a yeast WASP protein homologue.

KW Wiskott-Aldrich syndrome protein; WASP; actin cytoskeleton;

KW cell motility; actin polymerisation; cancer; parasite infection;

OS Schizosaccharomyces pombe.

PN WO200144292-A2.

PD 21-JUN-2001.

PF 15-DEC-2000; 2000WO-FR03569.

PR 16-DEC-1999; 99FR-0015900.

PA (CNRS) CENT NAT RECH SCI.

XX (CURRI-) INST CURRIE.

PI Noireaux V, Prost J, Sykes C, Friederich E, Golsteyn RM, Louvard D;

XX WPI: 2001-536241/59.
 DR N-PSDB; AAH77922.
 PT New fragments of WASP family proteins, useful for detecting and
 PT identifying modulators of actin cytoskeleton formation, potential
 PT anticancer and antiparasitic agents

Claim 14; Fig 8; 162pp; French.

CC The present sequence represents a WASP (Wiskott-Aldrich syndrome
 CC protein) homologue. Peptide fragments of WASP-family proteins of
 CC eukaryotic cells are used to prepare reagents for detecting compounds
 CC that inhibit or stimulate formation of the actin cytoskeleton, and
 CC thus inhibit or stimulate cell motility. The peptides are used to
 CC detect and identify compounds which are potentially useful for treating
 CC diseases associated with dysfunction of actin polymerisation,
 CC particularly metastatic cancer and parasite infection; as cytotoxic
 CC agents for inhibiting/stimulating formation of the actin cytoskeleton
 CC and for detecting side-effects, on actin polymerisation, of
 CC pharmaceuticals. By modulating actin polymerisation, these compounds
 CC affect cell motility, embryonic development, the immune response and
 CC wound repair.

SQ Sequence 574 AA;

Alignment Scores:

Pred. No.: 0.000153 Length: 574
 Score: 146.50 Matches: 53
 Percent Similarity: 39.41% Conservative: 14
 Best Local Similarity: 31.18% Mismatches: 64
 Query Match: 2.67% Indels: 39
 Gaps: 8

US-10-001-857-42 (1-3096) x AAG67370 (1-574)

```

QY 544 CTCCTGTCACCCGCGCGCGACCTGTCGAGACCGAATAAGACCGCCT 485
   |||||
Db 336 LeuProProProProProProProProProSerAlaAlaGlySerLeuProProPro 355
QY 484 CGG-----CCGCGCGCGCGCGCTCCCTCTCAGCCGCGACCGGACGCT 437
   |||||
Db 356 GlnGlyArgSerAlaProProProProProProProSerAla---ProSerThylArg 374
QY 436 GCACGATGCGTATGCGCGGACCGC-----GACCGCGCGCGCTCCCTA 389
   |||||
Db 375 GlnProProProProProProProProProSerAlaAlaValSerAsnProProAla 394
QY 388 CTCCTGCGGACCTTCGCGCTCCCGCGCACACC----- 356
   |||||
Db 395 IleProGlyArgSer---AlaProAlaLeuProProLeuGlyAsnAlaSerArgThSer 413
QY 355 -----CTCCCTGCGCGCGCGCTCCCTCTCGCGCGCA 326
   |||||
Db 414 ThrProProAlaProThrProProProSerLeuProProSerAlaProProSerLeuProPro 433
QY 325 TGGCGCGCGCTCTCTCTCTCGGAGATTAGACAGCATCGGACCGGAGTTATCTT 266
   |||||
Db 434 SerAlaProProSerLeu-----PrometGlyAlaProAlaAlaProProLeuProPro 451
QY 265 TTTCACCTCTGCGGAGTCTGCGTATGCGC-----GTTCC 227
   |||||
Db 452 SerAlaProProLeuProProProAlaGlyMetProAlaAlaProProLeuProPro 471
QY 226 CGGAAAGAGCGAGCGCGCGAGTCTCAGACCGCGCGCTGCGCTGCTGGAGAT 167
   |||||
Db 472 AlaAlaProAlaProProProAlaProAlaProAlaProAlaAla---ProValAlaSer 490
QY 166 TTCTTAGAAAGGTGACGCGGACCGGAGG 137
   |||||
Db 491 IleAlaGluLeuProGlnGlnAspGlyArg 500

```

RESULT 12

